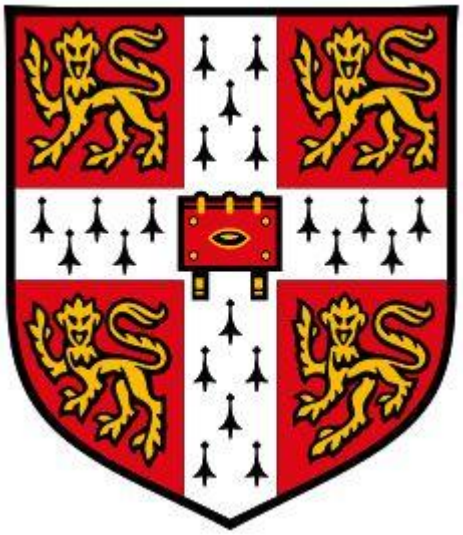
Synthesis and characterisation of peptide-based probes for quantitative multicolour STORM imaging



Edward John Robert Taylor

Downing College, Cambridge

This thesis is submitted for the degree of Doctor of Philosophy

September, 2017

Summary of synthesis and characterisation of peptide-based probes for quantitative multicolour STORM imaging by Edward Taylor

Current single molecule localisation microscopy methods allow for multicolour imaging of macromolecules in cells, and for a degree quantification on molecule numbers in one colour. However, that has not yet been an attempt to develop tools capable of quantitative imaging with multiple colours in cells.

This work addressed this challenge by designing linker peptides with chemospecific groups to allow attachment of activator and emitter dyes for STORM imaging, and a targeting module. The design ensured a stoichiometric ratio of targeting module to activator and emitter dyes.

Peptides with HaloTag ligands attached were labelled with various activator and emitter pairs and used to label HaloTag fusions of *S. pombe* and mouse embryonic stem cells. These peptides were found to bind non-specifically to various areas of both cell types, and did not localise to HaloTag protein, whereas controls did.

Another peptide was also labelled with activator-emitter pairs and attached to expressed anti-GFP and ant-mCherry nanobodies via native chemical ligation. The labelled anti-GFP nanobody was to demonstrate ensemble and single molecule imaging in *S. pombe*, as well as characterisation on single molecule surfaces in comparison to a conventional randomly labelled antibody. The stoichiometrically labelled nanobody had a more consistent number of photons detected per localisation, number of localisation per molecule and number of blinks per molecule, which implied that it could be more useful than randomly labelled nanobodies for counting experiments. It was also shown to be capable of specific laser activation for STORM imaging with both an Alexa405Cy5 and Cy3Cy5 pairs.

These anti-GFP and anti-mCherry nanobodies and peptide linker are new tools for both counting and multicolour imaging in super-resolution, which could be widely applied to constructs that are already tagged with GFP or mCherry.

Declaration

This thesis is the result of my own original work apart from Chapter 4, where the final data acquisition was carried out by David Virant, Ilijana Vojnivic, Bartosz Turkowyd from the Endesfelder group (Marburg, Germany) to correct for problems during original collection of the same data by myself. Also, mES cell culture and labelling was carried out by Dr Srinjan Basu.

This work is not similar to any that has been previously or is currently being submitted for any other degree, diploma or qualification at the University of Cambridge nor any other University or institution.

It does not exceed the prescribed word limit of 60,000.

Acknowledgements

First and foremost, I thank Professor Ernest Laue for the opportunity of studying in the unique environment that is Cambridge amongst so many distinguished scientists.

Secondly, every member of the Laue group who have always been happy to answer questions, suggest solutions and provide guidance. In particular, Dr Srinjan Basu and Dr David Lando who not only provided most day-to-day support but also spent days and possibly weeks reading this thesis. Also, our collaborators David Virant, Ilijana Vojnivic, Bartosz Turkowyd, and Dr Ulrike Endesfelder who stepped in with a mere nine months to go, directed and even carried out some of the microscopy elements of this work. Especially, Dr Ulrike Endesfelder who was willing to providing houseroom as well as spend hours of skyping time at strange times of night in those desperate last few months. She has been an essential knowledge resource for directing all the microscopy analysis. Thanks also to Dr Alex Herbert for incorporating the STORM analysis code into PeakFit – his expertise were absolutely essential.

On a personal level, thanks go to many members of the lab for making stressful times more relaxing over lunch, tea, coffee and the pub, as well as Dr Tim Stevens and Lisa-Maria Needham.

Downing College has been a great source of support and enjoyment over this time as well, both a financial benefit, for which I thank Dr Jay Stock, Jane Perks and Sarah Pickard amongst others, and as a great social scene. Some of the MCR members I particularly recognise as having contributed to making this time a hugely positive experience: Dr John Morgan, Ross Hubble, Jamie Fox, Dan Levy and Audrey Valreau.

Many other people and organisations have helped me enjoy this time as well including many of my old friends from Wadham, Downing College Boat Club, and Parkrun, without which these last few months would have likely been explosively stressful.

Of course, I also wish to thank my entire family – Taylors, Phillipses and Girards, not only for the love and support of nearly three decades, but the conversation, influence and education that have sculpted me as a person. Most importantly, my loving partner Emma, who has been steadfast and supportive for all this time, despite the infuriating person I can be. Thank you all so much.

Table of Contents

List of Figures

List of Tables

Table of Abbreviations

Ab	Antibody (full length $2V_H2V_L$)
Ac	Acetyl group
ACN	acetonitrile
Alexa488	AlexaFluor488
Alexa647	AlexaFluor647
Ala	alanine
Asp	Aspartate
BSA	Bovine Serum Albumin
BME	β-mercaptoethanol
C	Constant region of IgG heavy chain
C_H	Domain of constant region of IgG heavy chain
CCD	Charge coupled device
cDNA	Complementary DNA
CDR	Complementary determining region
CFP	Cyan fluorescent protein
Cnp1	*S. pombe* centromeric histone variant
COT	cyclooctatetraene
Cy3	Cyanine3
Cy5	Cyanine5
Da	Dalton
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DPBS	Dulbecco's phosphate buffered saline
*d*STORM	*Direct* stochastic optical reconstruction microscopy
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EMCCD	Electron Multiplying Charge Coupled Device
Fab	Fragment antigen binding (of an IgG)
FISH	Fluorescence *in situ* hybridisation
FP(s)	Fluorescent Protein(s)
fPALM	Fluorescence photoactivation localisation microscopy

Fc	Constant region of an IgG
FRET	Förster resonance energy transfer
GFP	Green Fluorescent Protein
GLOX	Glucose Oxidase
Glu	Glutamate
GSD	Ground-state depletion
H	IgG heavy chain
HaloTag	HaloTag protein
HaloTag ligand	Ligand for the HaloTag protein, which can be of the O2 or O4 variety.
Halo(O2)Peptide	Short peptide with an O2 N-terminal HaloTag ligand (see **Figure 2.1**)
Halo(O4)Peptide	Short peptide with an O4 N-terminal HaloTag ligand (see **Figure 2.1**)
hcAb	Heavy chain Ab (camelid)
HCl	Hydrochloric acid
HEMS	20 mM HEPES pH 7.5, 1 mM $MgCl_2$, 1 mM EDTA, 1 M sorbitol
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HIV-1	Human immunodeficiency virus type 1
His	histidine
HMSiR	Hydroxymethyl silicon rhodamine
ICS	Inter-system crossing
IF	Immunofluorescence
Ig	Immunoglobulin
IgG	Immunoglobulin G
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KDa	Kilo Dalton
k_{on}	on-rate of binding or dye blinking
L	IgG Light chain
LIF	Leukemia inhibitory factor
MALDI-MS	Matrix-assisted Laser Desorption Ionisation-Mass Spectroscopy
MEA	Mercaptoethylamine (also cysteamine)

mM	millimolar
mRNA	Messenger ribonucleic acid
mES cells	Mouse embryonic stem cells
MQ	MilliQ™ (Millipore)
N_3	Azide
Nb	Nanobody
NCL	Native Chemical Ligation
NHS	*N*-hydroxysuccinimide
NiNTA	Nickel-nitrilotriacetic acid
nm	nanometer
NPC	Nuclear pore complex
Oligo	oligonucleotide
PA	Photoactivatable (fluorophore)
PALM	Photoactivation localisation microscopy
PBS	Phosphate buffered saline
PC	Pair-correlation analysis
PCR	Polymerase chain reaction
PDB	Protein data bank
Pra	Propargylglycine
PS	Photoswitchable (fluorophore)
PSF	Point spread function
PVDF	Polyvinylidene fluoride
RESOLFT	reversible saturated optical fluorescence transitions
ROXS	reducing and oxidising systems
RP-HPLC	Reverse phase high pressure liquid chromatography
RNA-FISH	Ribonucleic acid-fluorescence *in situ* hybridisation
RNAPII	RNA polymerase II
ROI	Region of interest
RT-PCR	reverse transcription-polymerase chain reaction
SBz	S-benzyl group
SBz-peptide	Short peptide with a C-terminal S-benzyl group (see **Figure 1.18**)
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*

scFv	Single chain antibody
SDS-PAGE	Sodium dodecyl sulphate poly acrylamide gel electrophoresis
SiR	Silicon Rhodamine
SM	Single molecule
SMLM	Single-molecule localisation microscopy
SNR	Signal-to-noise ratio
SR	Super-resolution
SR-STORM	Spectrally resolved STORM
SSIM	saturated structured illumination microscopy
STED	stimulated emission depletion
STORM	Stochastic optical reconstruction microscop
Sulfo-Cy5	Sulfonated-Cyanine5
TBTA	Tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TEAA	Triethylammonium acetate
TES	30 mM Tris pH 8.0, 1 mM EDTA, 20% sucrose (w/v)
TEV	*Tobacco Etch Virus* nuclear-inclusion-a endopeptidase
TFA	Trifluoroacetic acid
TIRF	Total internal reflection fluorescence
TMR	Tetramethyl rhodamine
UV	Ultraviolet
V	Variable region of an Ab or derivative.
V_H	Variable fragment heavy chain of an IgG
V_HH	Variable fragment heavy chain of a camelid hcAb
V_L	Variable fragment light chain of an IgG
WL	White Light
YES	Yeast extract + supplements (media for growing *S. pombe*)

1 Introduction

One of the current challenges in biology is to understand how individual molecules such as proteins, nucleic acids, lipids and metabolites behave *in vivo*, with minimal external disturbance. This aim is necessary to better understand all the molecular mechanisms that underpin how life works and, therefore, medicine. A key method used to observe cells in their natural state is microscopy, which allows us to look at objects as small as 200 nanometers (nm) in diameter with minimal disturbance to the sample. When combined with biochemical methods modern microscopy is an essential tool for understanding how molecules, cells and small groups of cells such as embryos, colonies and biofilms work. This chapter focuses on the development of modern fluorescence and super-resolution microscopy together with the advantages and disadvantages of the field.

1.1 A historical perspective

1.1.1 White Light microscopy

White light microscopy was invented circa 1620 with several claims to the original microscope. These systems used two lenses, as shown in **Figure 1.1**. The first discoveries and reports focused on detailed descriptions and drawings of small creatures and objects such as insects, but soon led to the discovery of things that are too small for the naked eye to see, such as cells and single-celled organisms like bacteria.

The discovery of stains for different compounds, for example for bacterial cell walls, improved white light microscopy and led to the classification of gram-positive and gram-negative bacteria by their ability to stain with crystal violet. The construction of better lenses also allowed greater resolution to be achieved and more detail within cells to be seen. This work led to the detailed description of intracellular features such as the rough endoplasmic reticulum. The spots seen on this surface were later revealed to be ribosomes.

This type of microscopy led to significant advances and understanding of biology, but had limitations of resolution and density. When imaging with white light the only way to distinguish different materials is by their density, and since most of the intracellular mass is made up of equally dense aqueous solution and contains very small macromolecules there was only so much that could be observed with this method.

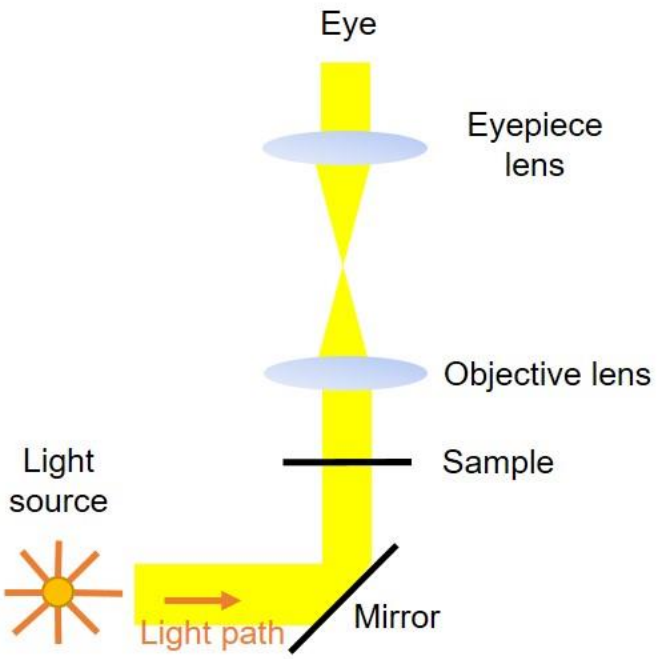


Figure 1.1 – Design of a white light microscope.

1.1.2 Fluorescence microscopy

Microscopy developed further with the discovery of fluorescent molecules in the nineteenth and early twentieth century. The original fluorescent compounds were naturally occurring molecules such as chlorophyll (Brewster 1834) and quinine (Herschel 1845), but soon fluorescein was synthesised (von Baeyer 1871). The principles of fluorescence were worked out in the early twentieth century as shown in **Figure 1.2**. In brief, light of a specific wavelength excites a molecule's electrons to a higher energy level, those electrons lose some energy through transitions between molecular vibrational and rotational energy levels, and when the electrons return to the ground state, light is re-emitted at a slightly longer wavelength. This increase in wavelength is referred to as the Stokes shift.

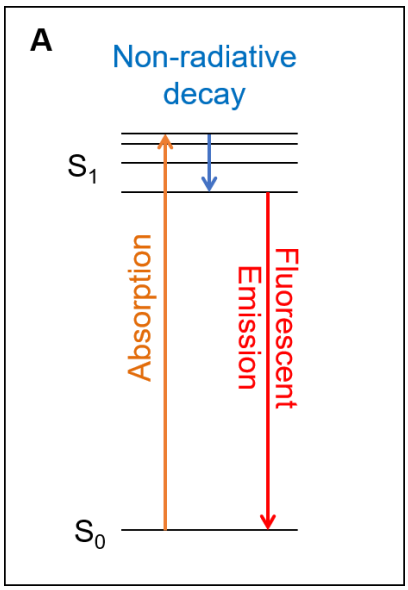


Figure 1.2 – Jablonski diagram of energy levels involved in fluorescence. S_0 is the singlet ground state, and S_1 are the singlet excited states.

The first fluorescence microscope was built by Heimstaedt and Lehmann (1911-1913), who used it to study cellular auto-fluorescence. The concept was also used to study dye-binding in live cells (Von Prowazek 1914). Immunofluorescence (IF) was then combined with the imaging technology by labelling of antibodies (Abs) with fluorescein isothiocyanate (FITC; Coons, Creech, and Jones 1941). Ever since, IF has been widely used to study target proteins, and FITC is still commonly used, although there are now many more dyes that can be used and Ab technology has improved dramatically (see **section 1.6.1**).

Modern fluorescence microscopes now use lasers instead of white light to specifically activate fluorophores at their excitation wavelength, and specific wavelength filters after the sample to remove the excitation laser light and maximise the fluorescence detected.

The advantages of IF are that the location of specific molecules, which are likely invisible to white light microscopes, can be seen within cells; Abs have high specificity and with different coloured fluorescent dyes attached to two different Abs, two different target molecules can be imaged in the same sample. This has increased the flexibility of microscopy and has enabled us to learn more about the location and actions of proteins in cells. IF has the disadvantage that it requires permeabilization to get Abs into cells, which can disturb the system.

Another limitation of fluorescence microscopy is that dyes like fluorescein cannot be resolved with molecular accuracy, due to the diffraction limit (Rayleigh 1879). This physical barrier considers all fluorophores (and other microscope and telescope specimens) as point sources, and a cone of light is collected from them by an objective lens. This produces an image of an Airy disc because the emitted light interferes with itself. Rayleigh somewhat arbitrarily decided that two point sources are optically resolvable if the central maximum of one Airy disc is no closer to the first diffraction minimum of the second, and thus derived a minimum resolvable distance called the Rayleigh Criterion, see **Equation 1.1**.

$$dr = \frac{0.61\lambda}{NA}$$

Equation 1.1 *– The Rayleigh Criterion*

Where *dr* is the resolution limit, *λ* is the wavelength of light emitted and *NA* is the numerical aperture of the microscope's lens system (NA = n.sinθ, where n is the refractive index of the medium in between the objective and sample, and θ is the half angle of the cone of light collected by the objective). With modern optics in oil immersion systems, the NA can be up to 1.6, which means that the resolution limit is 152.5-228.75 nm for visible light (400-600 nm). This is an obstacle when trying to resolve molecules that are very close together, because macromolecules such as proteins are often only ~3-4 nm in diameter. As a result, multiple small-to-average size proteins can be found within the diffraction limit, making quantitative measurements very difficult.

1.1.3 A spectrum of fluorescent proteins and dyes

Further advancements to fluorescence microscopy came with the discovery of genetically encoded fluorophores such as fluorescent proteins (FPs). The first fluorescent protein, green fluorescent protein (GFP), was discovered by Shimomura, Johnson and Saiga in 1962, and was cloned (Prasher et al. 1992), and later expressed in *E. coli* and *C. elegans* (Chalfie et al. 1994; Inouye and Tsuji 1994). Wild-type GFP was genetically modified in several laboratories, most notably the Tsien laboratory (1995) to improve photophysical characteristics, folding and stability as well as changing the absorption and emission wavelengths to produce a palette of fluorescent colours. This was extended by the discovery of a second wild-type FP – dsRed from the coral *Discosoma*, which also had mutants designed to modify characteristics and

colour. A selection of some FPs and dyes that are relevant to this chapter are shown, spectrally arranged in **Figure 1.3**.

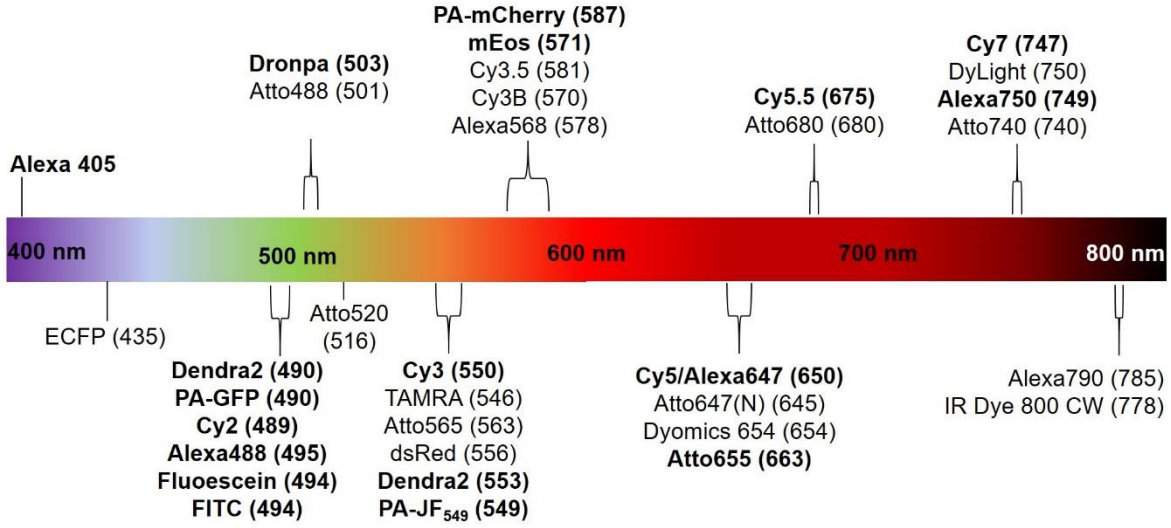


Figure 1.3 – Absorption maxima of some photoswitchable (PS) dyes and photoactivatable-FPs (PA-FPs) *from various sources (Dempsey et al. 2011; Grimm et al. 2016; Habuchi et al. 2006, Chroma® Spectra Viewer). Those in bold are referred to specifically in this thesis.*

FPs have the advantage over IF that they can be genetically tagged to a protein of interest and expressed in cells, which allows live cell imaging. The drawbacks of FPs are that they are often dimmer than fluorescent dyes used for IF; they often do not fold quickly so are not fluorescent and can induce artefacts such as multimerization of proteins caused by interactions of the FPs with each other.

Development of new families of fluorescent dyes including the cyanines, xanthenes (rhodamines) and oxazines, among others, have also improved the capabilities of fluorescence microscopy. Dyes are now available to stain nucleic acids, for example ethidium bromide and the Hoechst dyes; or lipids, e.g. Nile Red. Some are also sold commercially with reactive groups for labelling molecules of interest specifically, this can include Ab labelling, as discussed in **section 1.6.1**. The general advantages of modern dyes over FPs are their brightness, which makes detection and resolving them easier and the different types that bind different substances is practically very useful. The small size of dyes compared to FPs is also beneficial as is the fact that some of the families are net uncharged, so are membrane permeable, which suits live cell imaging studies. These advances now allow many different types of experiments to be done to yield important insight into the function of macromolecules in cells.

However, none of these advances in fluorophore technology can resolve fluorophores when they are closer together than the Rayleigh diffraction limit, so these types of

studies are termed ensemble or bulk experiments, where it is usually impossible to resolve single fluorophores. The limitation imposed by the diffraction of light means that certain types of information like numbers of molecules and precise locations cannot be attained with these types of fluorophores.

1.2 Super-resolution microscopy

Super-resolution (SR) microscopy, is a group of techniques that allow one to look at individual molecules rather than a population. They are instrumental in understanding sub-diffraction limit structures and interactions between macromolecules.

SR methods circumvent the Rayleigh diffraction limit for visible light (200-300 nm), allowing image resolution improvements of up to ten-fold (to ~20 nm). These methods include techniques such as: stimulated emission depletion (STED), ground-state depletion (GSD) and saturated structured illumination microscopy (SSIM), as well as inherently single-molecule localisation microscopy (SMLM) methods such as photoactivation localisation microscopy (PALM; Betzig et al, 2006; Hess et al, 2006), fluorescence PALM (fPALM), stochastic optical reconstruction microscopy (STORM; (Bates, Blosser, and Zhuang 2005; Rust, Bates, and Zhuang 2006)) and *direct* STORM (*d*STORM; Heilemann et al, 2008); for review see (Fernandez-Suarez & Ting, 2008). There is a general division between the SMLM methods and others that are considered as SR, although in some cases these SR experiments may actually be imaging single molecules. The difference in definition comes from the way these methods overcome the diffraction limit. SMLM methods exceed the diffraction limit by activating single fluorophores individually, whereas, other SR methods work on the principle of activating all the molecules within a certain limited volume, which if small enough may be a single molecule. Both categories and each individual method has distinct niches that suit them best, for example STED, as a scanning method, is generally used for fast collection of small fields of view or scanning of large or deep samples (Thorley, Pike, and Rappoport 2014), whereas, PALM is suited for live cell imaging, *d*STORM is most often used for imaging high resolution structure in fixed cells and STORM is used for simultaneously imaging many different colours. Here I focus on SMLM techniques.

1.3 Principles of single-molecule localisation microscopy (SMLM)

SMLM methods rely on the properties of photoactivatable (PA) and photoswitchable (PS) fluorophores, which allow imaging of one or a few molecules per camera exposure. The basic principle is that these fluorophores are normally invisible at the

imaging wavelength (in a dark state), but can be photoactivated/switched with a low power laser so that they fluoresce one-at-a-time, producing a single diffraction limited spot. A schematic depiction of localising a single fluorophore is shown in **Figure 1.4**. On a camera, the fluorescence of a single diffraction limited spot is detected across several pixels. Each of those pixels records a certain amount of light (photons) so that the pixel at the centre of the spot is brightest and the others around it less bright producing a histogram of light intensity across those pixels. The point source of this spot can then be identified by fitting the point spread function (PSF), typically a Gaussian curve, to the histogram of light intensity detected. Once a Gaussian has been fitted, it's centre point can be taken as the localisation of the molecule. The centre-point of the Gaussian cannot be calculated with absolute certainty, so the precision of the localisation is determined from the PSF. The better the PSF, the more precise (narrow) the localisation. The precision is therefore determined by how many photons are collected in the diffraction limited spot and how big it is in camera pixels. These are dependent on the brightness of the fluorophore, the length of the exposure needed to detect the spot, and the number of pixels across which that spot is spread.

An alternative method to using PA or PS fluorophores is to use fluorophores that diffuse in liquid of the sample, and transiently bind to a site of interest. These fluorophores are always in the on-state, so produce background fluorescence, but when they bind to a site the fluorescence produced at that spot is greater than background, which creates a diffraction limited spot, in the same way as a PA or PS fluorophore does, so they can be localised in the same way. This method is named point accumulation for imaging in nanoscale tomography (PAINT).

This process of activation/binding, imaging and localisation is carried out repeatedly to collect images of many molecules very quickly, and the images can then be super-imposed to produce a reconstructed image with many SM localisations. Localisation of each fluorophore to a sub-diffraction limited area means that the resolution of these images can be routinely be 20 nm (Rust, Bates, and Zhuang 2006; Bates et al. 2007; Cognet and Lounis 2014; Jones et al. 2011). Indeed, some imaging has been done to resolve fluorophores to much greater accuracy, such as 4Pi-SMLM in the Hell laboratory (Aquino et al. 2011), which has increased resolution to less than 10 nm by collection light from 2 objectives, and PAINT, which has been reported to achieve resolution of 5 nm (Schnitzbauer et al. 2017).

SMLM experiments therefore consist of: recording a video in which fluorophores are made to produce diffraction limited spots until there are all bleached. The reconstruction of a SR image from that video is then carried out by one of various programs to find all the localisations, and then plot them as desired. Several programs are openly available including rapidSTORM (Wolter et al. 2012, 2010); PeakFit (an ImageJ plugin; Palayret et al. 2015; Herbert 2014), GraspJ (Brede and Lakadamyali 2012), and ThunderSTORM (Ovensy et al, 2014) among many others. Comparisons and evaluations have been made of these different algorithms, which can act as a guide to which should be used for what purpose – seemingly there is no piece of software that is perfect for all experimental analyses (Small and Stahlheber 2014; Sage et al. 2015).

Whilst the use of SR and SMLM techniques has provided new insight into many areas of biology since their advent over a decade ago, there remain significant challenges to SMLM imaging methods, which derive from the basic concepts.

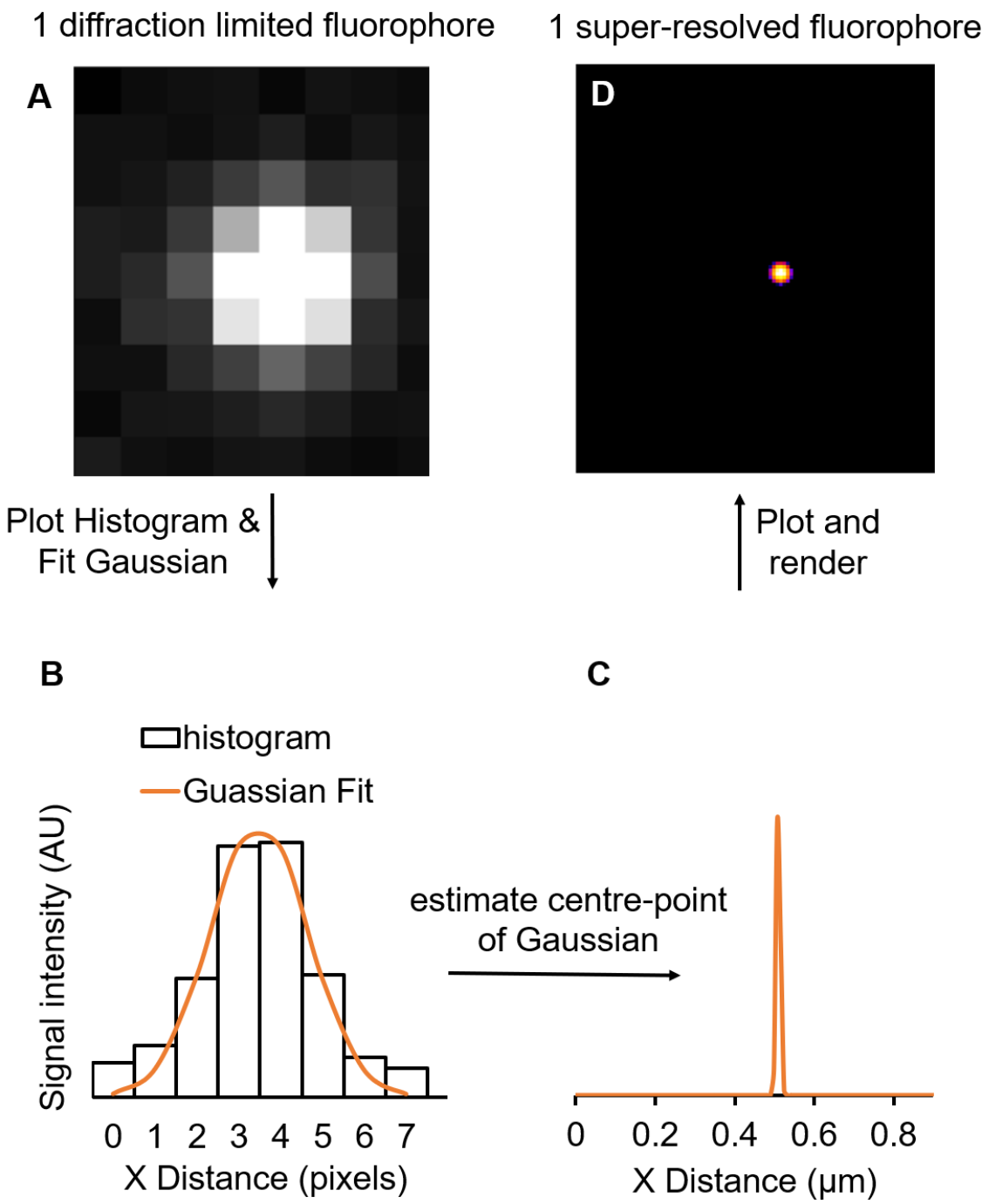


Figure 1.4 - Principles of SR imaging using a spot fitted by PeakFit (Herbert 2014). ***A*** *One diffraction limited spot is captured on a camera (pixel size 129 nm).* ***B*** *Representation of the histogram of number of counts detected in each pixel of the spot in the in the horizontal axis. A Gaussian is fit to the histogram.* ***C*** *The centre-point of the Gaussian is estimated mathematically.* ***D*** *The estimated precision with its associated error is used to create the image. This is repeated for every diffraction limited spot and combined to give a reconstructed image.*

1.3.1 STORM and *d*STORM

Two of the principle SMLM methods are STORM and *d*STORM, the difference between which is important to understand. Both are based upon the use of reversible saturated optical fluorescence transitions (RESOLFT) or reducing and oxidising systems (ROXS), and indeed often both use the same cyanine dyes; Alexa647, Cy5, Cy5.5 and Cy7. However, the key mechanisms are different, and in the case of STORM not entirely understood.

1.3.2 *d*STORM

Although STORM was introduced earlier (Bates, Blosser, and Zhuang 2005) than *d*STORM (Heilemann et al. 2008) it is more logical to consider the simpler mechanism of *d*STORM first.

*d*STORM was first carried out with the structurally analogous cyanine dyes Cy5 and Alexa647, which can be made to blink on and off for up to a hundred cycles under illumination with red laser light (Heilemann et al. 2008). A schematic for electron transitions in blinking dyes is shown in **Figure 1.5A**. Light excites the electrons in the ground state of the chromophore, causing it to fluoresce as normal. The rate of these transitions is on the nanosecond time scale, so in a typical frame on a camera, hundreds to thousands of cycles of this excitation and emission occur. In the presence of primary thiol reducing agents however, every few thousand excitation cycles, the excited electrons will instead transfer into a triplet state via inter-system crossing (ISC). This triplet state is not fluorescent. Cyanines can also be reversibly reduced on the unsaturated carbon bridge by primary thiols such as mercaptoethylamine (also known as cysteamine; MEA) or β-mercaptoethanol (BME), as shown in **Figure 1.5C**, which effectively quenches fluorescence. This ‘dark state’ exists on the millisecond time scale, until oxidation occurs. Once oxidised back to the ground state the chromophore is free to excite and fluoresce again. Hence, the reversible dark-state allows these dyes to blink, with most chromophores in the dark state at any one time at equilibrium. Importantly, this blinking is always carried out in reduced oxygen conditions by introducing an enzymatic oxygen scavenger system that slows oxidation back to the ground state.

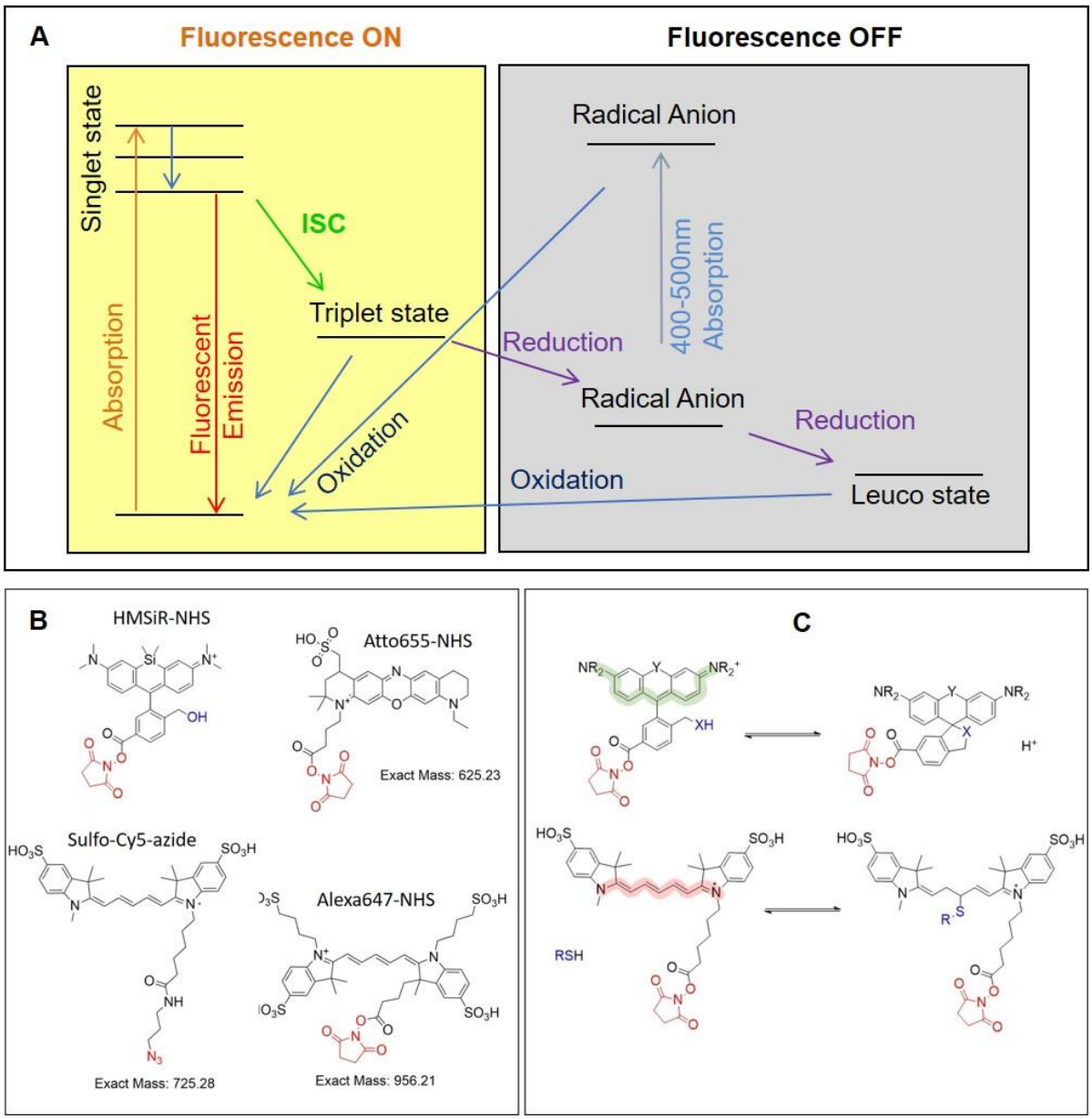


Figure 1.5 – dSTORM mechanisms and fluorophores. A *Energy level diagram for blinking fluorophores such as Cy5 and Alexa647, which can be repeated until chemical photobleaching (based upon figure 4 from Van De Linde et al. 2013).* ***B*** *The different dye classes available for dSTORM including the cyanines Alexa647 and Cy5, hydroxymethyl silicon rhodamine-N-hydroxy succinimide (HMSiR-NHS) and Atto655-NHS.* ***C*** *The chemical mechanisms that underlie photoswitching of cyanines and rhodamines (adapted from Turkowyd, Virant, and Endesfelder 2016; Uno et al. 2014). The highlighted blue groups are those involved in blinking, the green and red areas show the delocalized electron systems that produce fluorescence and the red groups can be involved in linking to molecules of interest.*

The reduced forms of these dyes have a different absorbance spectrum to the ground state, notably it is able to absorb ultraviolet (UV) and blue light, which increase the rate of oxidation (Heilemann et al. 2005, 2008). As a result, 405 nm and 488 nm lasers at low power are used to increase the on rate of cyanines to facilitate faster imaging.

*d*STORM has also been achieved with Atto655 and Atto680 dyes under specific buffer conditions of 2-10 mM MEA or glutathione in phosphate buffered saline (PBS; S. Van De Linde et al. 2008; Wilmes et al, 2013), rhodamines (Lukinavičius et al. 2013; Uno

et al. 2014), and various other dyes with different conditions (Dempsey et al. 2011). Only the rhodamine and cyanine blinking mechanisms have been studied in detail, via intramolecular cyclisation and reduction, respectively, as shown in **Figure 1.5C** (Uno et al. 2014; Turkowyd, Virant, and Endesfelder 2016). The cyclisation equilibria of rhodamines are responsive to pH and altered by the identity of multiple groups in the fluorophore (e.g. R and Y in **Figure 1.5**) and these observations were used to produce silicon rhodamine (SiR) derivatives such as hydroxymethyl silcon rhodamine (HMSiR) that blink more effectively, at lower laser intensity (Uno et al. 2014). They are particularly interesting because they have a net neutral charge and so are cell permeable, which is critical for live cell imaging studies.

Thus, *d*STORM is feasible with several different types of dyes, but is rarely reported with rhodamines or other non-cyanines – presumably because it is difficult to get the correct conditions for blinking.

1.3.3 STORM using activator-emitter pairs

STORM was developed by the Zhuang laboratory (Bates et al. 2012; Huang, Wang, et al. 2008; Dempsey et al. 2011, 2009; Bates, Blosser, and Zhuang 2005; Bates et al. 2007; Rust, Bates, and Zhuang 2006) and it has been reviewed by Van De Linde et al (2013). It uses activator-emitter pairs of dyes that rely on the same photoswitching mechanism of cyanines as with *d*STORM, see **Figure 1.5B** and **D**.

The activator fluorophore is excited by a laser, and the energy is transferred in a non-radiative manner to the emitter dye during its excitation/ISC/reduction cycle (S. Van De Linde et al. 2013), presumably via the radical anion, bringing it out of the dark state and back to the ground state. Whilst fluorescent, the emitter can be imaged at its appropriate wavelength, until it is either returned to the dark state by thiol reduction (Dempsey et al. 2009; S. Van De Linde et al. 2013) or photobleached. These activator-emitter pairs can be either covalently linked (Conley, Biteen, and Moerner 2008) or merely located within a short distance (~1-3 nm) from each other, for example on deoxyribonucleic acid (DNA) oligonucleotides (oligos) or a protein surface.

The main advantage of this approach is that an emitter can be paired with many different activators and the non-radiative energy transfer mechanism still functions because there is no requirement for spectral overlap of the main absorption/emission peaks, as with Förster resonance energy transfer (FRET). Hence, different activator-emitter pairs can be imaged in a sample simultaneously, allowing multicolour SMLM *in*

vitro (Bates et al. 2012, 2007) and in cells (Jones et al. 2011; Lakadamyali et al. 2012; Lubeck and Cai 2012; Huang, Jones, et al. 2008; Ricci et al. 2015). There is also the advantage that STORM acquisition only detects one read-out laser, whereas multicolour PALM and *d*STORM detect two colours which must be well aligned and chromatic aberration corrected for, which is not necessary for STORM. Furthermore, simultaneous data collection with two or more read-out lasers requires post-acquisition data analysis to align the channels, collected on separate areas of the camera. STORM does not require this step.

Thus far, mostly large or extracellular structures have been investigated using STORM (Conley, Biteen, and Moerner 2008; B. Huang et al. 2008; Lakadamyali et al. 2012; Löschberger et al. 2012; Xu, Babcock, and Zhuang 2012) and full advantage of STORM for imaging biological processes has not yet been taken, with some exceptions; (Dellagiacoma et al. 2010; Ricci et al. 2015). This is partly because means to efficiently make the various target probes with the appropriate dye pairs is lacking. This project will focus on the development of STORM for multicolour imaging of protein complexes by labelling each target molecule with specific activator-emitter pairs. It should thereby be possible to increase the number of proteins that can be imaged in a single sample, as well as potentially allowing quantification of numbers of target molecules. This has already been achieved when imaging messenger riobonucleic acids (mRNAs) using barcoding (Lubeck and Cai 2012), but target specificity and flexibility has not yet been maximised for proteins.

1.4 Applications of SMLM

SR techniques, with their key advantage of high spatial resolution are suited to experimental aims where other methods would not be as effective. High resolution is most often needed for structural imaging, multicolour imaging and quantitative imaging.

1.4.1 Structural imaging

The obvious benefit of SR imaging is the high spatial resolution, which allows biologists to look at previously unresolvable structures and observe nanoscale organisation of macromolecules. Many examples of this have been made in the last decade, with primary focus on cytoskeletal structures such as actin fibres (Malkusch et al. 2012), microtubules (S. Van De Linde et al. 2008; Heilemann et al. 2008; Bates et al. 2007; Lampe et al. 2012; Malkusch et al. 2012) and the nuclear pore complex (NPC; Pleiner et al. 2015; Löschberger et al. 2012) among many others. Thus, new methods are often compared using these benchmark systems.

1.4.2 Multicolour SMLM

One of the most potentially useful applications of SMLM is multicolour imaging, in which two or more different molecules are labelled with blinking fluorophores and imaged to give high resolution images of both molecules. This can be used in several different types of analysis for example co-localisation, which asks whether two molecules are close together in space? This can be strong evidence of direct or indirect interactions, which is important to understanding function in cells. Several different approaches have allowed SM multicolour imaging.

1.4.2.1 Multicolour *d*STORM

The principle of *d*STORM has been used in several different ways to achieve multicolour SM images. The conceptually simplest method is to use one blinking dye to label two targets sequentially, which has been done with Alexa647 (Nanguneri et al. 2012). This method relies upon complete bleaching of Alexa647 on one Ab, before labelling with a second. This is difficult to perform reliably because the first round of imaging is followed by a second staining, which may alter intracellular structures. The more commonly used method for dual colour *d*STORM is to use two spectrally separated dyes for different targets, so that the emission wavelengths can be separated by different emission filter sets. Fluorophores can then be imaged simultaneously or sequentially depending on requirements. The differences between the required imaging setups are shown in **Figure 1.6** and some examples are listed in

Table 1.1. Spectral de-mixing (SD-*d*STORM) of dyes imaged with the same laser has also been achieved, which works by illuminating two spectrally separated dyes with one laser, separating the emitted light with a dichroic mirror and recording both light paths. Localisations that appeared in both channels were assigned to a specific channel by intensity, which is characteristically different for Alexa647 and Alexa700, the latter being much dimmer.

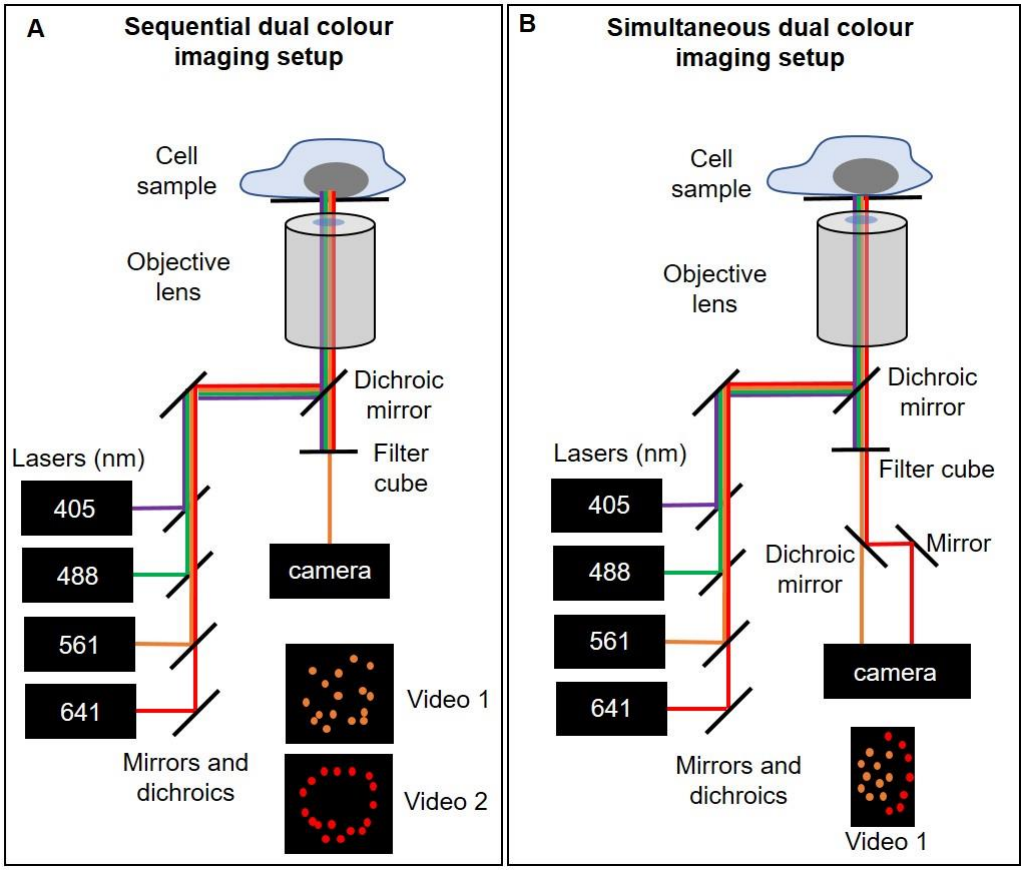


Figure 1.6 – Schematic imaging setup for sequential and simultaneous dual view dual colour imaging. A *Sequential imaging dual colour setup* ***B*** *Simultaneous two colour imaging setup with dual view. Both schemes are designed for orange (561 nm) and red (641 nm) imaging in this case, but could easily be altered for orange and green (488 nm) imaging by changing the filter sets in the filter cube. The 405 nm and 488 nm lasers could be used for photoactivation (e.g. mEos2 or Cy5).*

Dye 1	Dye 2	Area of study	Publication
Alexa647	Alexa488	Podosomes	(van den Dries et al. 2013)
Alexa647	Atto520	NPC	(Löschberger et al. 2012)
Alexa647	Alexa658	Alpha-synuclein	(Pinotsi et al. 2013)
Alexa647	Alexa655	Receptors, lipid rafts	(Gao et al. 2015)
Alexa647	Alexa700	Clathrin and tubulin	(Lampe et al. 2012)

***Table 1.1 – Dye combinations reported for multicolour d*STORM.**

Using a different optical and analytical system, Alexa647, Alexa680 and Alexa750 have also been imaged simultaneously to allow triple-colour *d*STORM in 2D and dual-colour in 3D, using the same imaging laser and determining fluorophore identity by the characteristic number of photons per SM blink with reportedly low crosstalk (Baddeley et al. 2011).

Clearly, Alexa647 is largely preferred for all variants of *d*STORM, which is due to its good photophysical properties and slow bleaching in oxygen scavenger systems used with thiols to facilitate blinking. Other blinking dyes like Alexa488, Alexa680, Alexa700, Alexa750 and Atto520 are either dimmer or photobleach more easily (van den Dries et al. 2013; Dempsey et al. 2011), which means that dual and triple colour *d*STORM is difficult although possible (Testa et al. 2010). The largest limitation for multicolour *d*STORM is the lack of good dyes that blink under the same buffer conditions and are resistant to photobleaching.

1.4.2.2 Multicolour PALM

PALM is another SMLM method that has been extensively exploited for use in multicolour experiments, using several different combinations of PA- and PS-FPs, derived from both the GFP, dsRed, EosFP and Dronpa families. Most FPs are imaged either in the green (488 nm) or the orange channel (561 nm), and so one FP emitting in each channel must be used. Dual colour PALM relies on selection of fluorophores and a scheme that minimises crosstalk between channels, as is the case with *d*STORM. The Betzig group used a method of pairing tandem dimer Eos (tdEos) and Dronpa, which relies on the RS nature of Dronpa to minimise its bleaching whilst activating both FPs with 405 nm light and imaging Eos at 561 nm to completion. They then serially activated Dronpa and imaged at 503 nm (Shroff et al. 2007). Other FP pairings have also been used, and are summarised in **Table 1.2**. (Brodehl et al. 2012).

Multicolour-PALM has the advantage of potential use in living cells because genetically encoded FPs no not require permeabilization. Another advantage is the FPs used are considerably smaller than Abs. The disadvantage is that sequential imaging will bleach some of the PS-CFP (cyan fluorescent protein) or Dronpa, so not all the second target can be imaged.

An alternative method for dual-colour PALM is simultaneous read-out of two spectrally distinct PA-FPs. This has been done using simultaneous imaging setups (**Figure 1.6**), with primarily green and red emitting PA-FP pairs, which are summarised in **Table 1.2**. The continual development of FPs means that optimal pairs for dual colour PALM are likely to be developed.

Green/ blue FP	Orange/ red FP	Acquisition method	Area of study	Publication
Dronpa RS-CFP(2)	tdEos EosFP	Sequential	Focal adhesion complexes, TCR clustering, cardiomyocyte and neuronal actin filaments, myogenesis	Shroff et al. 2007 Frost et al. 2010 Hsu and Baumgart 2011 Owen et al. 2010 Yao and Tjian 2011 Brodehl et al. 2012
PA-GFP	PA-mCherry PA-TagRFP	Simultaneous	Clathrin coated pits	Subach et al. 2009, 2010
rsKame RS-CFP2 Dendra2	PAmCherry	Simultaneous Sequential	Mitochondrial fission Clathrin coated pits Various systems	Rosenbloom et al. 2014 Annibale et al. 2012 Virant et al. 2017

Table 1.2 – FP pairings reported for multicolour PALM

The latest development of multicolour PALM imaging has been the application of primed conversion, which is a less toxic equivalent to 405 nm photoactivation. A combination of 488 nm and 750 nm light can be used with certain green-to-red FPs that have a threonine at position 69 (in Dendra2; Dempsey et al, 2015; Klementieva et al, 2016; Turkowyd et al, 2017). This was adapted for multicolour sequential imaging of PA-mCherry and Dendra2, because PA-mCherry is only UV-convertible whereas Dendra2 is primed convertible (Virant et al. 2017).

PALM, like *d*STORM, is, however, also limited in several ways: spectral separation is complicated because most PA and RS-FPs need a read-out laser as well as an activation laser, this makes schemes complicated and in particular too complicated to

allow three colour PALM. PALM has also been limited in the past by effects of multimerization caused by the FPs themselves. For example, wild-type dsRed naturally forms tetramers and so its derivatives such as mRFP (monomeric red fluorescent protein) and the mFruit series suffer from oligomerisation artefacts despite extensive efforts to prevent multimerization (Campbell et al. 2002; Shaner et al. 2004). The degree of oligomerisation of common PA and RS-FPs has been quantified (Cranfill et al. 2016), which shows that multimerization of monomerised FPs still occurs at high concentration.

1.4.2.3 Multicolour PALM-STORM

Whilst historically PALM and *d*STORM have been investigated separately, several reports have also combined the two techniques to produce multicolour images. This is feasible because there are some PA-FPs and blinking dyes that have compatible buffer conditions. The pair used most often is mEos2 and Alexa647 which are both able to activate and blink, respectively, in oxygen depleted hydrogen carbonate, pH 8, with 100 millimolar (mM) MEA to facilitate Alexa647 blinking. This pair has been imaged simultaneously using a dual view setup (Malkusch et al. 2012) and sequentially with Alexa647 read-out before Eos2 (Muranyi et al. 2013; Specht et al. 2013). These schemes were used in the Heilemann laboratory to study co-localisation with actin and its associated protein ezrin or human immunodeficiency virus type 1 (HIV-1) virion assembly, respectively.

A third study combined dual colour PALM of PA-GFP and PA-TagRFP with *d*STORM of HaloTag linked Atto655, resulting in triple colour super-resolution images in live cells (Wilmes et al. 2012). The scheme for this experiment was more like STORM than PALM or *d*STORM because an activation laser (405 nm) pulse was followed by several read-out frames using a 488 nm laser to detect PA-GFP, then several read-out frames using a 568 nm laser for PA-TagRFP, and several frames of 640 nm for Atto655. This meant the channel assignment of the localisations in the video were determined by what imaging laser was being used at the time that frame was recorded. This is one of only two reports of triple-colour imaging for PALM or *d*STORM (Baddeley et al. 2011), but neither has caught on as a standard approach, possibly because many localisations at wavelengths 1 and 2 will be lost when imaging at wavelength 3, this will result in approximately 1/3 of fluorophores of each type being imaged in total. This is critical for applications of multicolour SMLM that require localisations of as many of the present molecules as possible, like co-localisation and clustering.

1.4.2.4 Multicolour STORM

As described in **section 1.3.3**, STORM uses activator-emitter pairs which are excited by one laser and read-out by another. This method is ideally suited for multicolour imaging. The general scheme for all STORM experiments is the same, continuous read-out of the emitter dye and alternating pulses at wavelengths suiting the activator dye (usually 405 nm, 488 nm and 561 nm).

Two colour STORM has been carried using several different activator-emitter pairs, including Cy2Alexa647 and Cy3Alexa647 (Bates et al. 2007), Cy3/Alexa647 and Cy3/Alexa750 or Alexa405/Alexa647 and Cy3/Alexa647 in cells (Bates et al. 2012) and Alexa405/Alexa647 & Cy2/Alexa647 in to image neurones (Lakadamyali et al. 2012). Three colour imaging is also relatively straight forward by adding a third compatible activator dye and additional laser pulse into the STORM imaging scheme. It was initially carried out *in vitro* using Alexa405/AlexaA647, Cy2/Alexa647 & Cy3/Alexa647 (Bates et al. 2012), and then used to image neurones with the same three dye pairs (Lakadamyali et al. 2012), as well as other pair combinations Alexa405/Cy5, Alexa488/Cy5 and Alexa555/Cy5 in 3D (Huang, Wang, et al. 2008). The structures of these activators are shown in **Figure 1.7**, except for Alexa555, for which the structure is not published.

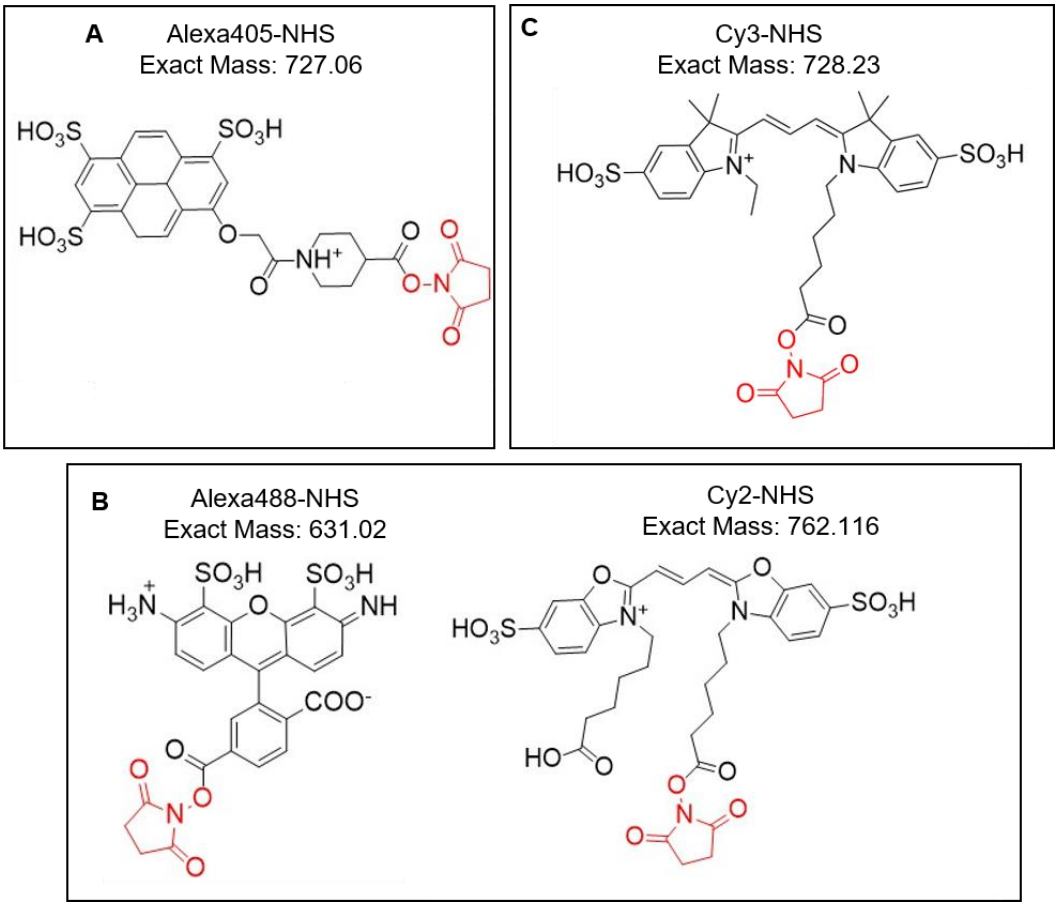


Figure 1.7 – Activator dyes available for STORM imaging. A *405 nm activated,* ***B*** *488 nm activated,* ***C*** *561 nm activated.*

However, to proceed beyond three activators is difficult because each activator dye must be spectrally separated from the others in order to allow activation by a specific laser, as well as read-out of the emitter by a fourth laser. To overcome this, the Zhuang laboratory combined different activators with two different emitters, Alexa647 and Alexa750, which are themselves spectrally separated enough to be imaged specifically by lasers at ~650 nm and 750 nm. This facilitated six-colour imaging *in vitro* with Alexa405/Alexa647, Cy2/Alexa647, Cy3/Alexa647, Alexa405/Alexa750, Cy2/Alexa750 and Cy3/Alexa750 pairs (Bates et al. 2012). This second STORM emitter system has only once been used for imaging in a cellular context, when 9 activator-emitter pairs were made using Alexa405, Alexa488 and Cy3 activators and Cy5, Alexa680 and Alexa750 emitters (Lubeck and Cai 2012). The pairs were arranged using short 25-mer DNA oligos for ribonucleic acid-fluorescence *in situ* hybridisation (RNA-FISH). Combining the 9 pairs into spatially arranged barcodes allowed detection of 32 mRNAs of interest. The authors imaged all 32 different barcodes although only 9 distinct colours, in the conventional STORM sense were present. This work yielded

quantifiable amounts of each mRNA through spatial and STORM barcoding as a read-out of cellular stress in response to calcium, but did generate reconstructed images.

Clearly STORM has more 'colours' available than *d*STORM and PALM, which is a key strength. One of the limitations, however, is that all these reports use Abs or DNA probes to target the STORM pairs. This prevents the use of STORM in live cell imaging as these labelling methods require cellular permeabilization. Furthermore, the complex probabilistic algorithms for channel assignment are not openly available and this limits the widespread use of STORM. Thus, only a handful of laboratories have reported STORM imaging.

1.4.2.5 Multicolour PAINTing

Point accumulation for imaging in nanoscale topography (PAINT) and its derivatives are another set of methods that have allowed multicolour SR imaging. It takes advantage of non-blinking fluorophores in solution that bind a target and become immobilised to produce a diffraction limited spot, before dissociating and the spot disappearing. If the concentration of fluorophores is carefully controlled and the association and dissociation rates are within a certain range, then fluorophores will bind and dissociate from solution to produce a blinking behaviour that can be super-resolved in the same way as blinking dyes or PA-FPs. This method was first demonstrated using Nile Red, which binds to membranes (Sharonov and Hochstrasser 2006), or membrane proteins with Cy5 or ATTO647N labelled Abs (Giannone et al. 2010), leading to multicolour 'SPIRAPAINT' images (Lew et al. 2011).

PAINT was combined with DNA-origami (Rothemund 2006) to allow Atto655-labelled short oligos (imager strands) to be used in 'DNA-PAINT' (Jungmann et al. 2010). Different imager strands with the same dye could then be combined with Abs linked to DNA docking strands to stain fixed cells. Sequential washes allowed 'dual-colour' images with only one dye, where the two colours are distinguished temporally and by imager strand sequence specificity for a different docking strand on a different Ab (Jungmann et al. 2010). The same multicolour concept was used to image 10 different colours *in vitro* using DNA-origami, four-colour images in 2D and three-colour images in 3D using single dyes (either Atto655 or Cy3B) in fixed HeLa cells. The method was dubbed exchange-PAINT (Jungmann et al. 2014).

The advantages of multicolour PAINT compared to the previously used methods are clear in that only a single dye and imaging laser combined with different Ab-docking

strand constructs are necessary. This removes the barrier of spectral overlap and channel crosstalk. The drawbacks are that the Ab-docking strand reagents take considerable design and purification, and are not compatible with live cell imaging. Furthermore, PAINT has been reported to cause cell swelling during long acquisition times (Legant et al. 2016), which is difficult to correct for and may perturb samples. These disadvantages and the fact that many groups do not have experience with DNA-origami for method development and optimisation may be the reason that exchange-PAINT has not been widely used by others, with very few exceptions (Zanacchi et al. 2017).

1.4.2.6 Spectrally resolved STORM (SR-STORM)

SR-STORM (Z. Zhang et al. 2015) uses similar wavelength absorbing dyes that can be made to blink, but which cannot be separated by emission filter sets, which pass light within ~40nm. The dyes can be distinguished instead based on small differences in their emission spectra, by measuring an intensity weighted mean emission wavelength in a separate beam path, as little as 2.5 nm. This requires a different setup to conventional PALM, *d*STORM and STORM microscopes, where emitted light is collected by two objective lenses. The light from path one is passed to the camera to give spatial information, but light from the second path two is passed through a prism, to refract the light before collection on a separate area of the camera. This means that each dot has spatial and spectral information, so it can be super-resolved and the spectral mean taken, which can assign a localisation to dyes with very similar emission spectra with <2% crosstalk, superior to that of STORM. This system has the benefit of only requiring a single laser for imaging, so little optical alignment, but the complication of a second optical path containing the prism.

SR-STORM was initially used to image 4 colours in 2D and 2 colours in 3D for biological samples (Z. Zhang et al. 2015) and has since been used to image NileRed PAINT binding to different intracellular membranes, each of which shifts its spectral mean differently (Moon et al. 2017). Further multicolour imaging has yet to be reported, but seems likely given that this optical setup removes the spectral limitations of multicolour *d*STORM imaging. This method seems likely to challenge both STORM and PAINT as one of the most useful multicolour SMLM methods.

1.4.3 Quantitation

Another benefit of SMLM is that observing single fluorophores may allow counting the precise number of molecules, which is useful for discerning macromolecule function, for example stoichiometry of proteins in complexes and multimerization states.

1.4.3.1 Diffraction limited quantification

Quantification has been attempted using diffraction limited fluorescence, using purely fluorescence intensity and stepwise photobleaching to count molecules with some success (Leake et al. 2006; Durisic et al. 2012; Ulbrich & Isacoff, 2007) and even with diffraction limited imaging of PA-FPs like Dendra2 (Specht et al. 2013). This method relies on a single fluorophore emitting a roughly constant intensity and when it is bleached a discernible intensity being lost, as depicted for a diffraction limited spot with a stoichiometry of 2 in **Figure 1.8**. One example was the use of GFP photobleaching when connected to a bacterial flagellar protein MotB (Leake et al. 2006). This study produced good evidence of MotB-GFP stoichiometry in the motor and in kinetics of MotB turnover.

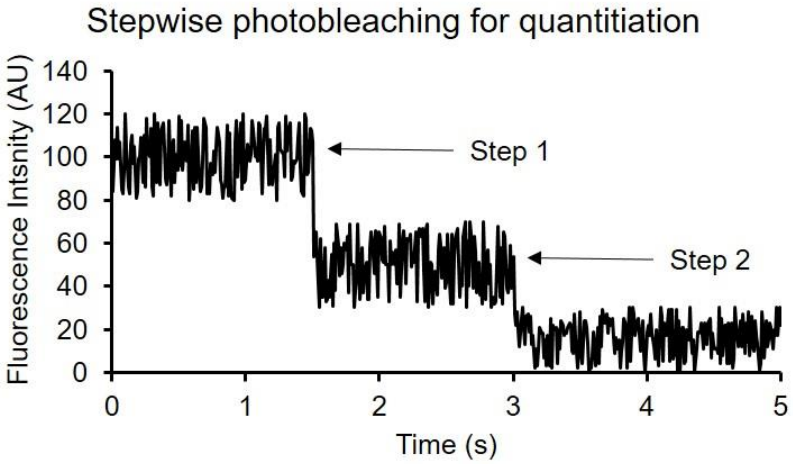


Figure 1.8 – Stepwise photobleaching can be used for quantitation. *An Example of fluorescence intensity detected for a particular diffraction limited spot, each step down in intensity indicates the photobleaching of one fluorophore.*

These diffraction-limited methods of quantitation can be useful in some cases, but are less effective when high density, high background fluorescence and movement is encountered because intensity is continuously distributed, and only an average intensity per fluorophore can be measured, with significant error in most cases. Errors in the measurement of fluorescence intensity can be improved by using total internal reflection fluorescence (TIRF) illumination, which reduces background, but this limits

measurement to surfaces close to the coverslip, so cannot be used for nuclear or cytoplasmic proteins in thick cells.

SMLM techniques provide the advantage that they count a discrete number of localisations and these can be grouped spatially and temporally into ‘tracks’ to allow counting of how many blinks per molecule are achieved. This has been used as a more accurate measure of protein numbers, but is not without is pitfalls.

1.4.3.2 Quantitation with PALM

PALM and *d*STORM using tdEos or mEos2 with Cy5 or DyLight 648, respectively, have both been used in the Gaus laboratory to look at clustering of signalling proteins such as Lat, Lck and Src in the T-cell membrane (Owen et al. 2010; Williamson et al. 2011) and to study HIV replication and assembly (Pereira et al. 2012). These studies all used Ripley’s k-function to measure and plot clustering of localisations at the membrane before and after T-cell stimulation. These studies showed that SR could visualise microdomains of associated proteins on the scale of 100-200 nm, which was not possible with ensemble TIRF. This clustering, however, whilst semi-quantitative was not counting absolute numbers of molecules because the studies could not account for multiple localisations of each fluorophore due to blinking, a common problem in SMLM quantitation. Blinking is a larger problem in *d*STORM where fluorophores like Cy5 can blink up to a hundred times (Heilemann et al. 2005; Dempsey et al. 2011), but has also been shown to be a problem with commonly used PA-FPs for PALM, like mEos2 (Annibale et al. 2010; McKinney et al. 2009).

Several reports have attempted to consider and correct for various sources of error in counting. Lando et al (2012), measured the number of blinks per mEos2 molecule using spatially separated fluorophores within an *S. pombe* cell, and calculated an average of 1.98 blinks per molecule. To calculate the number of mEos2-Cnp1 molecules present in the dense centromeric focus, the number of localisations was divided by 2. More complex methods were not appropriate due to the high density of Cnp1 in *S. pombe.* Another study quantified overcounting due to blinking and undercounting due to localisations being incorrectly grouped together by spatial proximity (Lee et al. 2012). This work used a kinetic model for blinking and measurements of the respective rate constants, which could be used in analysis for minimising both over and undercounting – ‘spatiotemporal clustering’. It was determined that Dendra2 is better suited to counting than mEos2 because it blinks less

owing to a slower transition rate in the dark state, corroborating previous evidence (Annibale et al. 2011). Several different methods to account for spatiotemporal clustering have been published, one used a data-optimised distance and time threshold for each localisation to group localisations into molecules (Coltharp, Kessler, and Xiao 2012). An alternative method to account for blinking as a source of uncertainty was developed by Sengupta et al (2011; Sengupta and Lippincott-Schwartz 2012). They used pair-correlation analysis (PC) to analytically calculate the probability that each localisation belongs to a given fluorophore through space and time. The algorithm increases the likelihood of correctly assigning localisations to the right molecule, but works less well at high fluorophore density and longer dark times between blinks, so it is more suitable to PALM than *d*STORM or STORM. Both the Coltharp and Sengupta methods have been implemented in the PeakFit open source software for SMLM data analysis (Herbert 2014). Work on PC also noted that PS-CFP blinks less than Dendra2, and if it could be made to only blink once, analysis would be far more certain about its conclusions. PC analysis has been extended to RS probes and immuno electron microscopy (Veatch et al. 2012).

A third method to account for blinking in quantitation experiments was also developed using a similar kinetic model to Lee et al (2012) using mEos2 (Hummer, Fricke, and Heilemann 2016). The kinetic scheme was investigated analytically, and knowledge of several variables was necessary to find the number of mEos2 molecules within a diffraction limited spot. By simulating results, fixing one parameter at a time and varying the others, the best fit for the modelled number of fluorophores to the real data could be determined. Using multimeric DNA-origami constructs the model was shown to be very accurate.

Another source of error for counting with PALM is that not all PA-FPs are detectable, presumably due to incomplete maturation (Durisic et al. 2012). This maturation problem has been quantified for mEos2 in *S. cerevisiase* by measuring known numbers of PA-FPs, and it was found that ~40% of mEos2 molecules are undetectable (Puchner et al. 2013), but it is likely that this value as well as blinking and photophysical characteristics will vary extensively based on cellular, sub-cellular and even local environment conditions.

1.4.3.3 Quantification using *d*STORM

Quantification was initially far simpler using PALM than *d*STORM, simply because PA-FPs blink less than the preferred organic dyes like Alexa647 and Cy5 (Durisic et al. 2012). However, some dyes can be made to blink very sparsely, and so spatiotemporal clustering analysis (Lee et al. 2012) can be applied in a similar way. This has been carried for tetramethyl rhodamine (TMR) and SiR in *d*STORM analysis of RNA polymerase II (RNAPII; Zhao et al. 2014).

The analytical model developed for mEos2 in the Heilemann laboratory (Hummer, Fricke, and Heilemann 2016), was also used in cells (Krüger et al. 2017) and then developed to consider the differences between PALM and *d*STORM, notably that *d*STORM experiments start with the fluorophores on, rather than off as in PALM. This resulted in a similar analysis that could also count numbers of Alexa647 molecules within a diffraction limited spot on DNA-Origami structures *in vitro* (Karathanasis et al. 2017). This work is particularly promising given the wide spread use of Alexa647, but has yet to be used in cells for biological work.

1.4.3.4 Quantification using STORM

STORM has also seen relatively little use for quantitation, presumably because the required emitter dyes all blink a lot more than PA-FPs and rhodamines, which complicates spatiotemporal analysis. One paper from the Lakadamyali laboratory reported quantification of nucleosomes and RNAPII in mouse embryonic stem cells (mES cells, Ricci et al. 2015), by calibrating the number of emitter localisations to the known number of nucleosomes in designed arrays. They did not count absolute numbers, but produced median numbers of nucleosomes per cluster, which was then comparable between cells. This allowed them to study the number of localisations of RNAPII in dual colour STORM, which showed that pluripotent stem cells have fewer nucleosomes per clutch than more differentiated neural progenitor cells. Perhaps corresponding with formation of more heterochromatin during differentiation. Since then the same laboratory has extended quantification with STORM by using DNA-origami as a calibration tool that allowed them to count protein constituents at the NPC (Zanacchi et al. 2017). This method, potentially combined with that of Karathanasis et al (2017), are promising developments for quantitative STORM.

1.4.3.5 Quantification using PAINT

PAINT has also been developed for quantitative use (qPAINT) by relating the number of available binding sites to the dark time, bright time, influx rate of imager strand and binding on-rate (k_{on}) of binding for a particular system (Jungmann et al. 2016). The dark time can be measured in a control experiment, and the influx is a controlled factor. The approach was benchmarked by counting visible DNA-origami spots *in vitro*, NPC protein Nup98 in U2OS cells and Brp in *Drosophila*, with validation via comparison with previous literature.

This is an interesting approach to quantitative imaging because the fluorophores used (Cy3b or Atto655) do not blink under these imaging conditions, so the over- and undercounting effects of blinking are not a problem, and neither is bleaching, because any one dye molecule is irrelevant. These are significant advantages over counting with PALM, *d*STORM or STORM, but there are also drawbacks. Firstly, there is a practical drawback that controls require DNA-origami, which is only used in few laboratories, and secondly, quantification of proteins is carried out with Abs covalently labelled with docking strands. Ab labelling itself has been highlighted as an under-counting risk, as not all sites might be bound.

1.4.3.6 Summary of quantitation using SMLM

SMLM clearly has several advantages for counting molecules compared to lower resolution methods, but the methodology comes with specific problems, many of which have been identified and considered.

Overcounting is caused by over-labelling e.g. blinking of dyes and FPs. Under-counting can occur due to under-labelling with Abs, FP slow maturation and transient transfections in mammalian cells.

One important aim is to label targets in a one-to-one ratio or as close as possible, which is difficult with large, bivalent targeting modules like Abs (Durisic, Cuervo, and Lakadamyali 2014) and to carefully account for FP maturation.

Spatiotemporal problems of blinking and localisation uncertainty can often be solved with PC analysis (Sengupta and Lippincott-Schwartz 2012; Sengupta et al. 2011) or spatiotemporal clustering (Lee et al. 2012; Zhao et al. 2014) with the availability of good control samples to measure fluorophore parameters. However, these methods are not valid with fluorophores that blink extensively or in very densely labelled

samples, hence fluorophores that blink less have been preferred like Dendra2 and PS-CFP. A very recent development of spatiotemporal clustering has been applied to the PA-FP mEos2 (Hummer, Fricke, and Heilemann 2016) and the *d*STORM dye Alexa647 *in vitro* (Karathanasis et al. 2017), but this has only been applied to cellular imaging for mEos2 (Krüger et al. 2017).

In cases where absolute quantification is not possible, relative numbers can still be of use for certain biological questions, such as clustering using Ripley's K-function (Owen et al. 2010; Williamson et al. 2011; Pereira et al. 2012) and calibration (Ricci et al. 2015). Further development of the algorithms and techniques will undoubtedly lead to better counting ability, although the method must be carefully linked to the question and the likely data patterns.

1.4.3.7 Quantification and targeting

Whilst fluorophores and their characteristics such as brightness and blinking are critical to quantification, the method used for fluorophore targeting is also an important factor. FPs are genetically tagged, which is desirable for counting proteins, as it allows one-to-one labelling so that number of fluorophores equals number of target proteins. However, the FP mEos2 has been shown to be invisible in 40% of cases (Puchner et al. 2013; Durisic et al. 2012), and this is likely the case for other FPs. Given their low brightness compared to dyes, FPs may not be ideal for quantification.

Dyes also have their problems for counting target proteins because they cannot be genetically encoded, and so they must be targeted with other approaches, each of which has its own strengths and weaknesses. These are discussed in detail in **section 1.6**.

1.5 Advantages and disadvantages of SMLM techniques

SMLM has several advantages over ensemble imaging and other SR techniques, but each SM experiment has specific advantages and disadvantages to consider. It is important to understand these for each method and each biological question to achieve meaningful results.

1.5.1 PALM

PALM has the strength of genetically tagged FPs that allow live cell experiments with up to two colours. Its weaknesses include the FPs that are all dimmer than the widely used *d*STORM and STORM dyes Alexa647 and Cy5, with lower extinction coefficients and quantum yields, which can limit spatial precision. FPs are also known to multimerise, fold incompletely and blink which can all cause artefacts in quantification experiments although progress with algorithms may offset these.

1.5.2 *d*STORM

This method has the advantage that the dyes of choice are Alexa647 and Cy5, are brighter than FPs, which can lead to better spatial resolution than PALM in structural imaging experiments, which have been done with up to two colours, although more in some specific cases such as SR-STORM.

Whilst labelling cells with Alexa647 and Cy5 is incompatible with live cell imaging, other dyes such as rhodamines like Atto655 are compatible with live cell *d*STORM experiments (Wombacher et al. 2010; Wilmes et al. 2012; Klein et al. 2011).

Also, the targeting approach can cause artefacts including for quantification experiments as discussed in **section 1.6**. Quantification is complicated further by the blinking of the dyes, although progress has been made with recent modelling approaches (Karathanasis et al. 2017). Finally, most dyes are not as bright or blink less than Alexa647 and Cy5, and may have incompatible blinking conditions, which makes multicolour *d*STORM difficult.

1.5.3 STORM

STORM has similar strengths and weaknesses to *d*STORM since it uses the same cyanine dyes, however it has the key advantage that it is far more suited to multicolour imaging, with up to 12 colours imaged *in vitro*. Its use in quantitative imaging has been minimal due to the artefacts caused by fluorophore targeting and blinking, although

some of these can now be accounted for (Karathanasis et al. 2017; Zanacchi et al. 2017)**.**

1.5.4 PAINT

PAINT has also been developed for multicolour SMLM with up to ten colours used in fixed cells, and has also been extended to quantitative imaging with qPAINT. The main advantage in applying PAINT is that the method in unaffected fluorophore bleaching, because new probes bind, this can allow very high resolution (Schnitzbauer et al. 2017), and cover all target sites that may be missed by bleaching. The lack of requirement for blinking dyes may also be an advantage, since more choice is available. These advantages have made PAINT a technique that is growing in usage.

The main weakness of PAINT is the requirement for dye targeting, which may cause similar artefacts to STORM and *d*STORM and prevents live cell imaging. It has also been reported that exchange-PAINT causes cell swelling during the long wash sequences which may also produce artefacts.

In summary, none of the SMLM techniques are universally suited to all biological questions and applications, so it is essential to carefully consider which approach is most suited to each experiment.

1.6 Fluorophore targeting

SMLM of biological targets requires fluorophores to be targeted to the molecule of interest, usually protein or nucleic acid, as specifically as possible, with minimum disturbance of the system. This is relatively straight forward using genetically coded FPs to N- or C-terminally tag the gene of interest. However, targeting of *d*STORM and STORM dyes to proteins of interest is not as straight forward, with different methods having advantages and disadvantages.

1.6.1 Antibodies (Abs) and their derivatives

Abs and their recombinant derivatives are one of the commonly used methods for targeting fluorophores to proteins of interest. It is therefore important to understand how these different proteins are made, and how the design and synthesis can affect SMLM applications, like STORM.

1.6.1.1 Conventional immunoglobulins

Conventional mammalian immunoglobulins (Ig) are divided into five classes: IgM, IgG, IgA, IgD and IgE (these can vary between families and between vertebrates). The most commonly used for biochemical purposes are IgGs, which are generated via immunisation with a target antigen. A polyclonal antiserum can then be obtained from the immunised individual by purifying Abs from blood serum, which contains many different IgGs, each binding different epitopes on the antigen. Each of these IgGs is produced by B cells from a different lineage.

Monoclonal IgGs (mAbs) are therefore usually used in SMLM protein targeting. They are obtained by immunising rabbits or mice spleen with an antigen, and then be fusing them with myeloma cells to form hybridomas, which are immortalised (Köhler and Milstein 1975). Individual colonies grown from single hybridoma cells each produce a different mAb. These clones can be selected for antigen affinity and those with highest affinity can be purified and compared for cross-reactivity using enzyme-linked immunosorbent assay (ELISA) screens. They can even have their affinity increased by phage display (Smith 1985), as reviewed by Clackson et al (1991).

The structure of Igs has important indications for to understanding the benefits of different Abs, see Hanly, Artwohl and Bennett et al (1995). Briefly, Each Ig molecule consist of 4 chains, two heavy chains (H) and two light (L) chains, the L and H chains interact through non-covalent interactions of hydrophobic residues, and the two H chains are then also held together by disulphide bonds as shown in **Figure 1.9**. Each

of the four chains has a variable (V) and a constant (C) region. The variable regions of the H and L chains pair up such that two V_H-V_L pairs are formed, each containing the epitope binding site – the paratope. The C regions of the longer H chains also interact to form the constant region (Fc), which is constant between Igs of the same class, and imparts biological function such as macrophage stimulation. The H chain of mammalian IgGs contains four domains: the V_H, and 3 domains in the constant region (C_H1-3) Thus, each IgG contains two paratopes and an Fc, the latter of which is not always useful for biochemical applications.

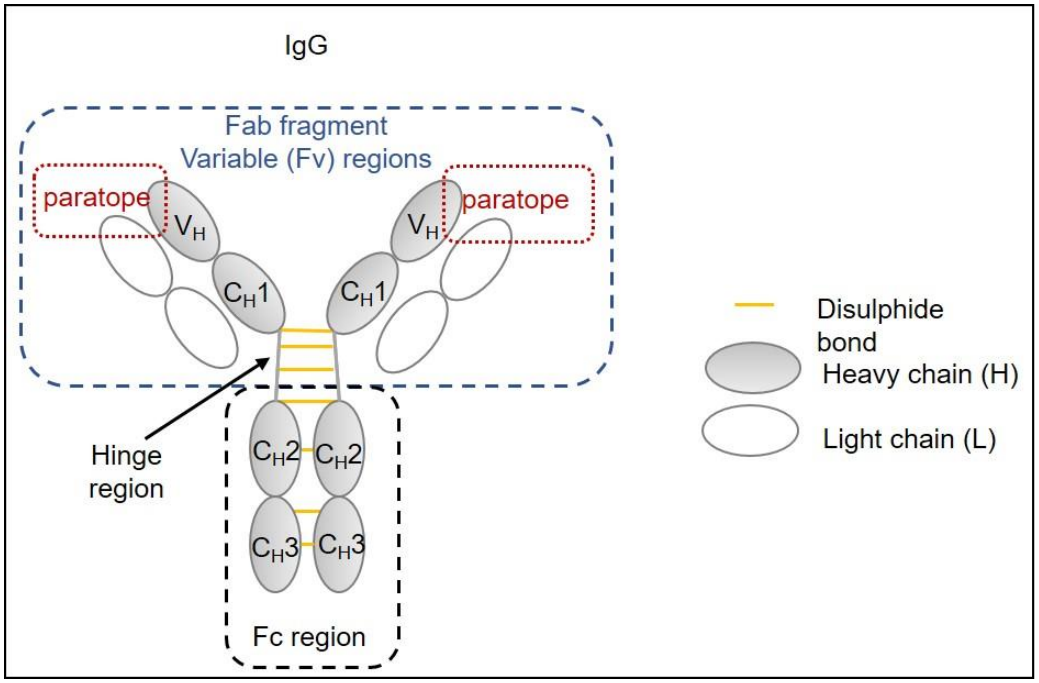


Figure 1.9 – Summary of generalised Ab structure and its derivatives. *Generalised conventional IgG Ab structure containing indications of derivative fragment antigen binding (Fab) in blue. Each domain of the heavy chain is labelled.*

1.6.1.2 Fragment antigen binding (Fabs)

Since the Fc of an IgG is sometimes not useful or counterproductive, this region can be removed by proteolytic cleavage of an IgG (usually using pepsin or papain) leaving just the Fabs (**Figure 1.10**). These Fabs can then be purified and used in biochemical contexts, such as IF (Stasevich et al. 2014). They have the advantage over IgGs of being smaller, which may allow them to penetrate denser targets in cells more effectively for IF; they are more structurally ordered than IgGs, and without the Fc region, are unlikely to induce a host response when used as treatments.

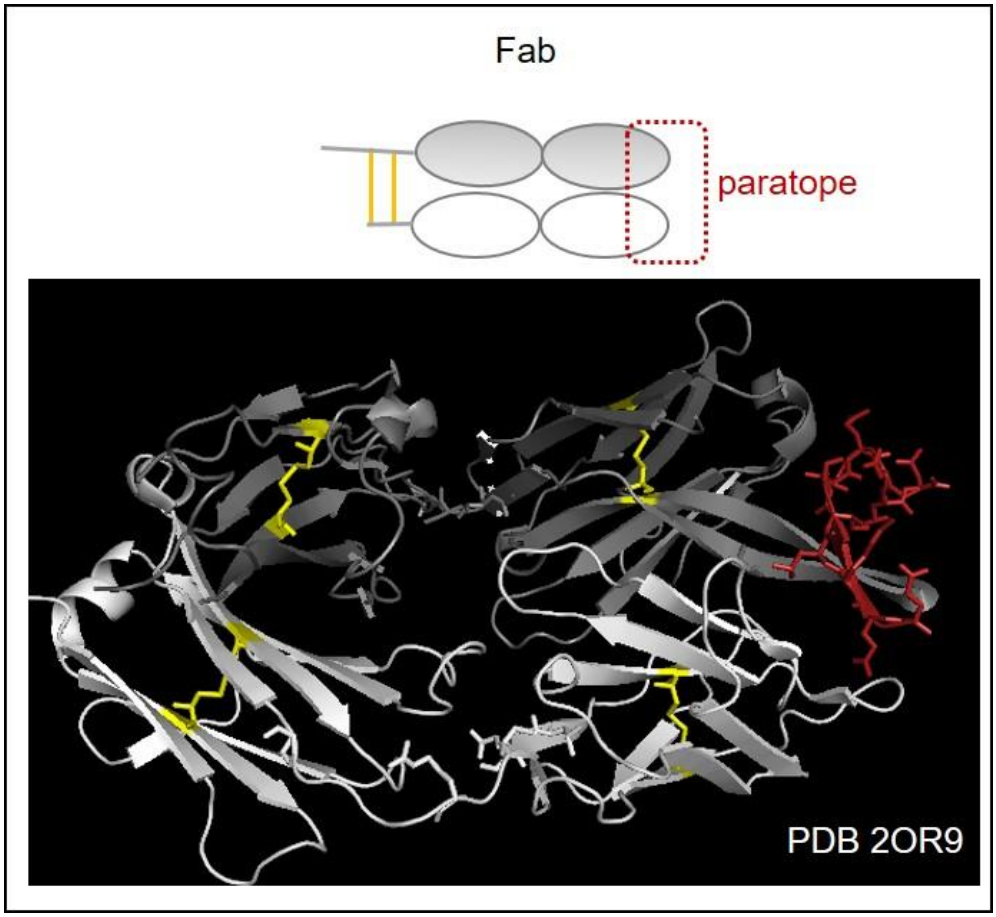


Figure 1.10 – Generalised structure of a Fab, *with a crystal structure of an example, an anti-c-Myc Fab, adapted from* Krauß et al (2008*; protein data bank (PDB): 2OR9). The epitope is shown in red, the H chain in grey and L chain in white, with anchoring disulphides in yellow.*

1.6.1.3 Single-chained variable fragments (scFvs)

scFvs are made by linking the V_H and V_L domains via a linker, usually in a V_H-linker-V_L arrangement, by various methods including polymerase chain reaction (PCR) assembly (reviewed in Ahmad et al. 2012). This construct can be expressed in bacteria or other systems, the generalised structure and an example are shown in **Figure 1.11**.

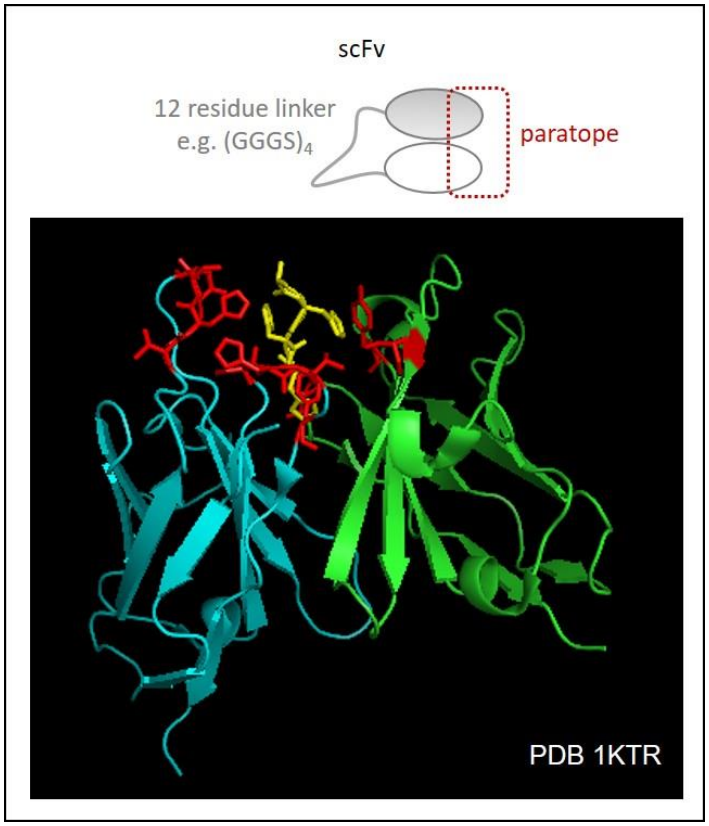


Figure 1.11 – General scFv structure with an example structure of the anti-his scFv described by Kaufman et al (2002;PDB 1KTR). *Green and blue highlight the V_H and V_L domains of the scFv and in yellow is the tetra-his peptide epitope, with the paratope residues in red. The linker is disordered and so not resolved in the structure, as are three residues at the N-terminus.*

The advantages of scFvs include their small size, the lack of protease treatment needed during purification, and they can be expressed from non-hybridoma systems such as *Escherichia coli* (*E. coli*; Jung and Pluckthun 1997), yeast (Boder and Wittrup 1997), mammalian and insect systems (A. B. H. Choo et al. 2002), which are all simpler and produce greater amounts of protein more quickly. To date many scFvs have been produced. However, design of correctly folding and expressing scFvs is not trivial particularly because hydrophobic residues are exposed on the V_H and V_L domains. They have not been used in SR or SMLM, due to their laborious and low yield expression, and their similar performance to conventional Abs (Fornasiero and Opazo 2015).

1.6.1.4 Camelidae and shark Abs

It was discovered in 1993 that camelidae have an intrinsically different Ab (Hamers-Casterman et al. 1993). As shown in the schematic in **Figure 1.12A** they only contain two identical heavy chains, hence they are referred to as heavy chain Abs (hcAbs). In addition, the region equivalent to the Fab in conventional IgGs contains only one discrete domain compared to the 2 domains per chain in IgGs, due to the absence of constant region 1 (C_H1). These hcAbs form the same basic "Y" shape as an IgG and function in a similar way, with C_H2 and C_H3 forming a constant region for biological function. Convergent hcAbs have also been found in sharks (Greenberg et al. 1995) and are suspected in other species, with the selection force thought to be that the small variable domain (V_HH) size allows better binding to some antigens, particularly the active clefts of antigenic enzymes (Flajnik, Deschacht, and Muyldermans 2011).

The missing C_H1 domain from hcAbs means that their digestion with papain or equivalent forms only a single domain V_HH, as shown in **Figure 1.12B** – there is no Fab equivalent for this type of Ab. These V_HHs are ~15 Kilo Daltons (KDa), the smallest antigen-binding immunoglobulin fragment yet discovered, and were first expressed in Muyldermans laboratory (Arbabi Ghahroudi et al. 1997).

The presence of only one protein chain means that the process of deriving Ab binding fragments is more straightforward than with conventional Abs. They are obtained by immunising a camelid such as a llama, collecting mRNA from the blood or spleen and synthesising complementary DNA (cDNA) via RT-PCR (reverse transcription-PCR). V_HH genes are obtained by PCR amplifying the relevant regions of the coding sequence and ligating them into plasmids. The product of this is a nanobody (Nb).

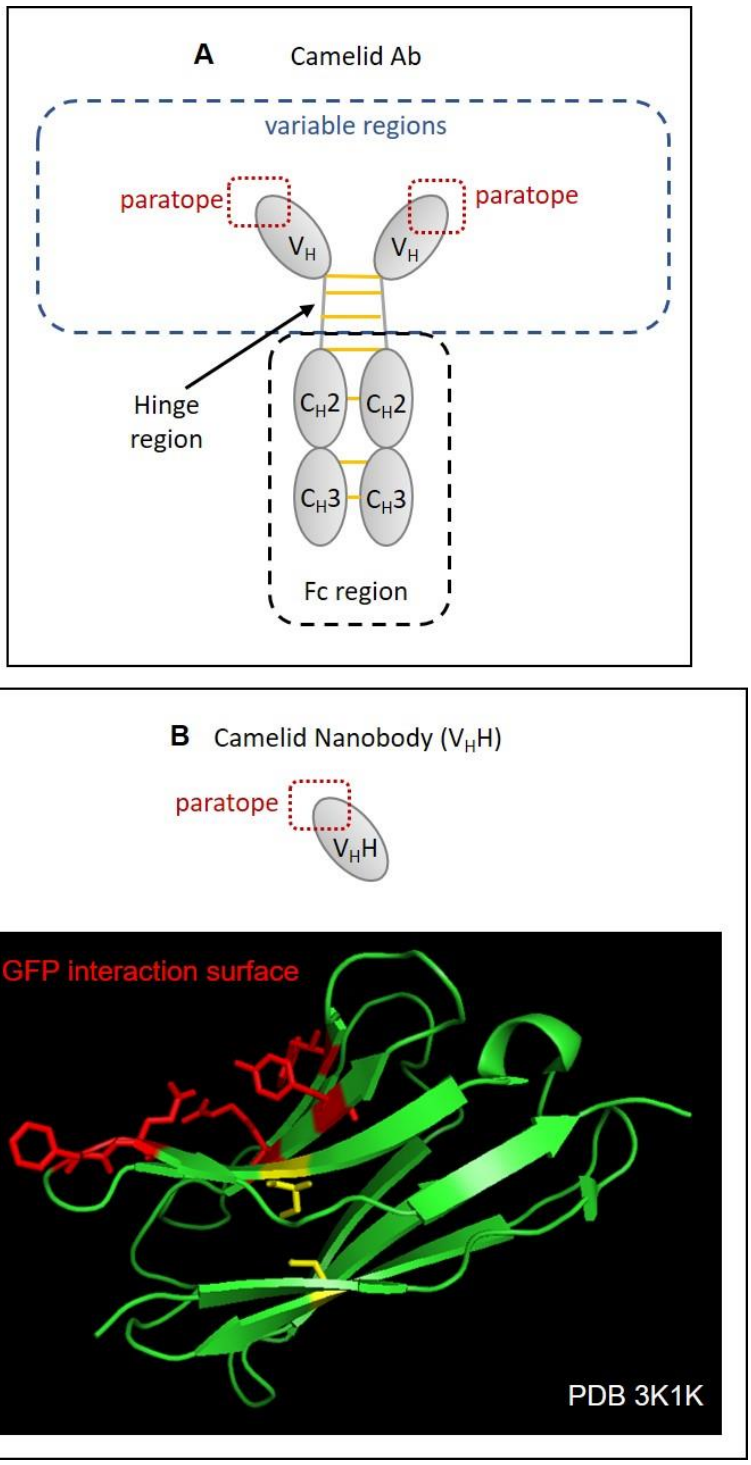


Figure 1.12 – Summary of camelid Ab and derived Nb structure. A *Generalised structure of a camelid Ab.* ***B*** *Generalised structure of a Nb, with a crystal structure of an example; anti-GFP nanobody (PDB:3K1K, Kubala et al, 2010). The paratope is shown in red, with anchoring disulphide in yellow.*

1.6.1.5 Nanobodies

In the short time since Nbs have been developed as potential tools, it has been evident that they have several advantageous properties.

Firstly, their small size (~15 KDa compared to scFvs ~28 KDa) has various advantages. It is useful for IF because they are more capable of penetrating dense molecular structures like nuclear chromatin, thus there is currently a considerable amount of effort going into developing Nbs that target useful epitopes.

A second advantage is that they have greater stability, with similar affinity to mAbs. A direct comparison of mouse mAbs and Nbs showed that mono-valent Nbs had similar specificity (in the nanomolar range), whilst most Nbs also had greatly increased heat stability – maintaining antigen binding up to 90ºC, and target affinity in up to 50% ethanol or 0.25 M ammonium thiocyanate (van Der Linden et al. 1999). This is corroborated by evidence that Nbs are stable when kept at 37ºC for one week, compared to tested scFvs that lost all binding activity within 60 hours (Arbabi Ghahroudi et al. 1997). The increased stability may also make Nbs more amenable to in house production from bacterial cells without unfolding. The stability is partly due to a single disulphide bond anchoring together the two β-sheets, which is not solvent accessible and prevents passive reduction. For SMLM it is particularly important to have the most stable targeting proteins, because the STORM buffers can be relatively harsh, involving high concentrations of additives like cyclooctatetraene (COT), 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox), MEA and BME, which could all damage extracellular proteins. Fluorescence *in situ* hybridisation (FISH) experiments also require maximum stability because the conditions used for these protocols usually involve heating and formamide, which may also cause damage before fixation.

Thirdly, of the few Nbs that have been produced and investigated, several have been targeted to common genetic tags such as GFP (Kubala et al. 2010; Kirchhofer et al. 2010; Fridy et al. 2014) and mCherry (Fridy et al. 2014). Binding of common epitopes like these could be potentially useful for IF protocols because raising Nbs against each individual protein of interest would take large amounts of time and may yield Nbs with too poor affinity. A drawback of the tag method however, is that the tag itself may cause artefacts in the cellular context such as preventing normal protein-protein interactions or inducing multimerization (Gahlmann and Moerner 2013).

Nbs have also been shown to be useful for IF. Several Nbs raised directly against different NPC components; Nup85, 93, 98, 155, and the Nup62Nup58Nup54 complex. These Nbs were linked to Alexa647 and used to image the NPC in *Xenopus* XL177 cells and showed excellent resolution (Pleiner et al. 2015). This was particularly notable because the Nbs were labelled in a one-to-one fashion using maleimide linkages for *d*STORM imaging. This example highlighted not only the utility of Nbs but also the effect of using one-to-one linking to *d*STORM dyes.

1.6.1.6 Labelling immune derived proteins is Poisson-distibuted

Most IF experiments have used labelled Abs to target them to a protein of interest (Ricci et al. 2015; Lakadamyali et al. 2012; Chaumeil et al. 2006, 2008). This is good for high sensitivity but suboptimal for some aspects of SMLM such as quantification of target molecules because labelling of Abs using amine-reactive dyes yields random labelling, which has Poisson distribution as shown in **Figure 1.13A-B** (Grunwaldt 2002; Sako, Minoghchi, and Yanagida 2000; Schnaible and Przybylski 1999; Vira et al. 2010). This means that even if the number of fluorophores can be accurately and precisely measured, it may not yield the correct number of target proteins, because each Ab could be labelled with between zero and five dyes. This effect is magnified with dual colour experiments to study co-localisation (as shown in **Figure 1.13C)**, because assuming an average number of dyes per molecule of 0.7-1, which is typical, only ~37-50% of each interaction partner is labelled. For the case of one dye per Ab on average, only ~ 43% of interacting pairs will be theoretically visible, so more than half of the interactions are lost. Only ~16% of all interactions will be labelled with both Abs having one dye attached, so quantification of co-localisation is very difficult, especially when the target proteins are already in low copy-number. This is important because defining protein stoichiometry in complexes in cells is of considerable interest.

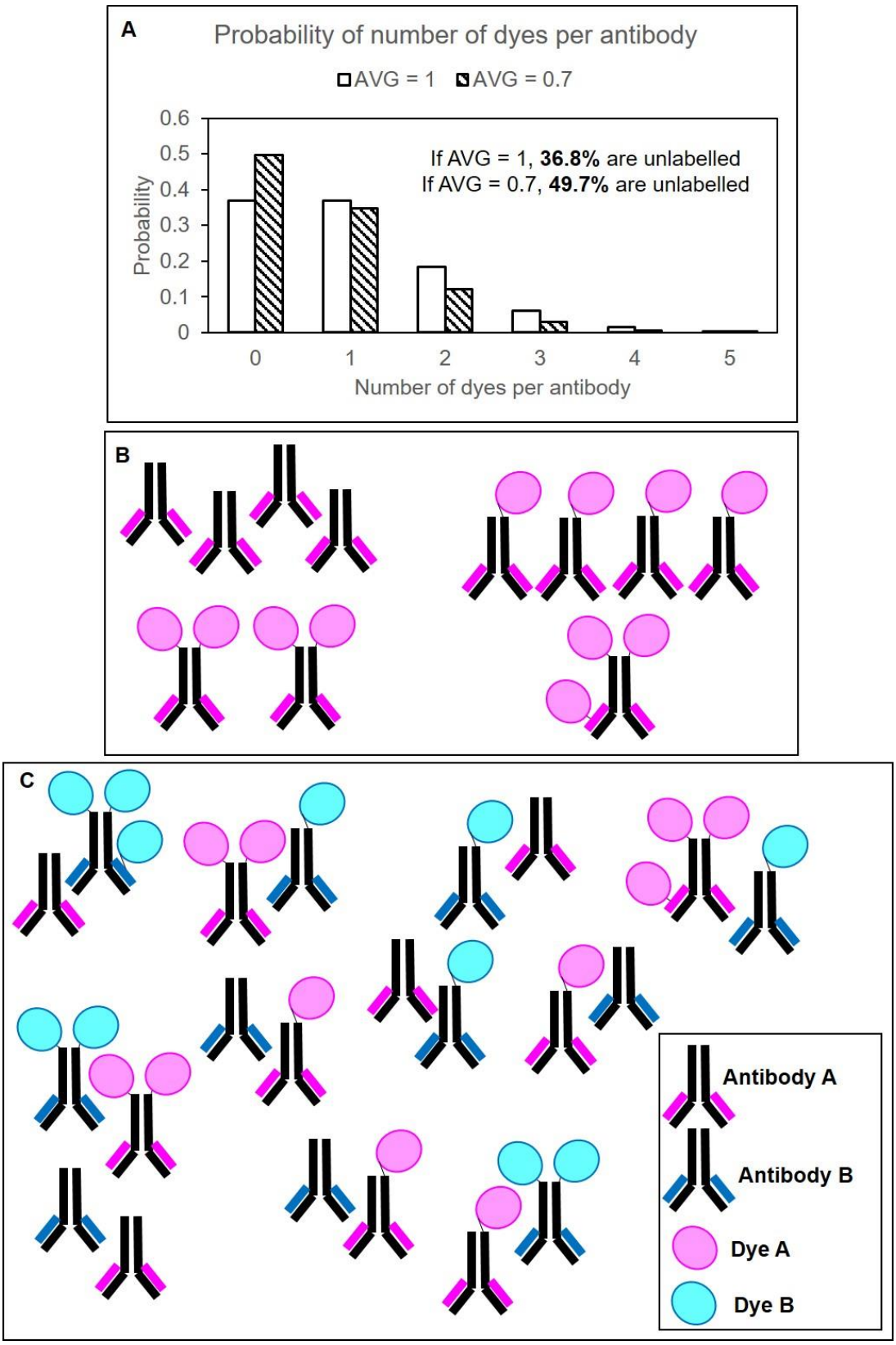


Figure 1.13 – Labelling Abs with dyes has Poisson distribution. A *Poisson distribution for an average of 1 and 0.7 dyes per Ab.* ***B*** *How Poisson distributed labelling affects the number of fluorophores per Ab with an average labelling of 1.* ***C*** *Pairing up of two labelled Abs A and B, both labelled with on average 1 dye per Ab, to represent co-localization of proteins A and B. On average in ~64% of cases one or both Abs are unlabeled, so the co-localization is invisible. On average in 16% of cases both Abs A and B singly labelled for quantification.*

1.6.2 Dehalogenases and HaloTag

Another modern method of linking fluorophores to proteins of interest is the HaloTag protein. This system was developed by Los et al (2008), from a *Rhodococcus* dehalogenase named DhaA. This class of protein removes halogen atoms from alkane chains is part of the α/β hydrolyase family of enzymes which also include acetylcholinesterase, dienelactone hydrolase, lipase and carboxypeptidase II (Ollis et al. 1992). This family contains a conserved catalytic triad structure that allows water hydrolysis of their respective ligands but each divergent enzyme has a differently structured binding cavity to accommodate the different substrate. The triad consists of two nucleophiles: cysteine, aspartate or glutamate (Cys, Asp or Glu) and a conserved histidine (His). The nucleophile forms a covalent intermediate, and then base catalysis by the His regenerates the nucleophile and releases the product (Franken et al. 1991).

The dehalogenases can bind a variety of 1-haloalkanes, however, the most common substrate appears to be 1,2-dichloroethane. One of the best studied organisms is *Xanthobacter autotrophicus*, which can survive on 1,2-dichloroethane as a sole carbon and energy source (Franken et al. 1991), in which DhaA catalyses the first step in the degradation.

The substitution mutation H289Q of the *Xanthobacter* DhaA enzyme was catalytically inactive and produced a covalent product between the alkane and the nucleophilic aspartate residue of the catalytic triad (Pries et al. 1995). A *Rhodococcus* homolog (Newman et al. 1999) had a significantly deeper and wider binding pocket providing a greater flexibility in the substrates it could bind. This was developed for specific binding of the HaloTag ligand in the presence of large modifying groups like fluorophores at low concentrations suitable for *in vivo* labelling, with low levels of cross-reactivity (Los et al. 2008). The crystal structure of a similar *Rhodococcus* gene product with the HaloTag ligand is shown in **Figure 1.14A** and **B**, as adapted from a protein data bank (PDB) structure by Liu et al (2017; PDB 5VNP). It shows the surface cavity and how the HaloTag ligand docks into that cavity.

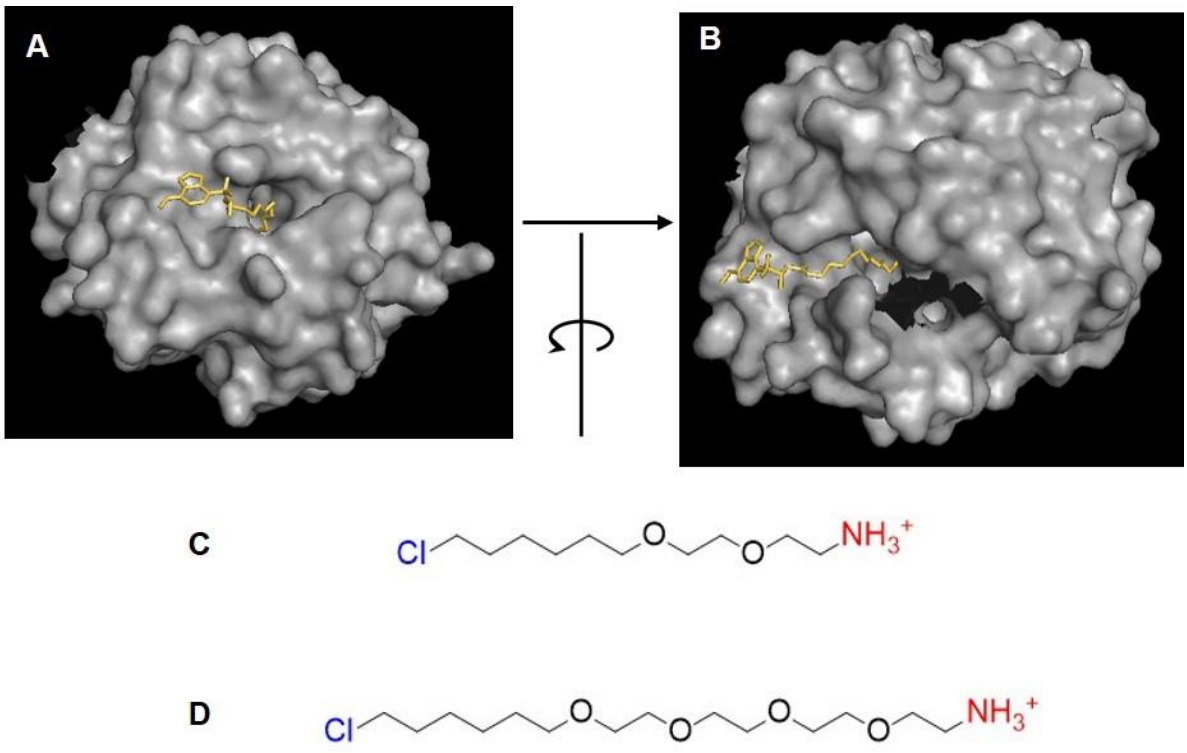


Figure 1.14 – HaloTag and its ligand structures. ***A*** *Crystal structure of a DhaA mutant together with its ligand (a sulfonyl-benzoxadiazole moiety with a sarcosine linker between this and the halo(O2) ligand; PDB5VNP; Liu et al. 2017). The image shows the ligand moiety on the surface and the HaloTag ligand region entering a cavity in the protein. The ligand and attached moiety are shown in yellow, the protein surface is shown in grey.* ***B*** *The same image rotated 90º, with several resides hidden to reveal the halo ligand biding cavity beneath the protein surface.* ***C*** *HaloTag(O2) ligand and* ***D*** *HaloTag(O4) ligand both with the chlorine (Cl) atom still attached, which is removed upon covalent linkage.*

The key advantage of this labelling method for fluorescence is the ability to genetically tag endogenous proteins. Then label them with any of a variety of fluorescent dyes, whereas genetically encoding proteins with FP tags limits experiments to one colour. For a different colour, significant time must be taken to clone in a different FP. The requisite disadvantage is that fluorescent dyes attached to the ligand must pass through the cell membrane (and the cell wall in the case of micro-organisms) to label the HaloTag protein. This is problematic for live cell and fixed cell experiments using organisms that have complex cell wall structures. Although the HaloTag ligands alone can pass through cell membranes, only a limited number of uncharged fluorescent dyes such as TMR and Atto655 can. Whilst these dyes have been reportedly used for *d*STORM (Grimm et al. 2015, 2016; Wilmes et al. 2012; Uno et al. 2014), STORM has only been reported for the cyanine dyes Alexa647 and Cy5, which each have multiple charges and so are largely unable to pass through cell membranes directly.

Commercially available HaloTag ligands include those with pre-linked fluorophores produced by Promega, however, many more dyes can be made HaloTag compatible by combining the HaloTag ligands with compatible reactive groups such as NHS esters

attached to dyes, which are readily available for most dyes. For this purpose, there are two different commercial versions of the HaloTag ligand; the Halo(O2) ligand and the Halo(O4) ligand. The difference between the two ligands is that the Halo(O2) ligand, which shown chemically in **Figure 1.14C**, contains 2 oxygen atoms in its alkyl linker, whereas the Halo(O4) ligand in **Figure 1.14D** contains 4 oxygen atoms. This means that the Halo(O4) ligand is longer and may provide higher affinity in cases where the dye may sterically hinder docking and binding of the Halo(O2) ligand. In most published cases the Halo(O2) ligand is used (Wilmes et al. 2012; Y. Zhang et al. 2006; So, Yao, and Rao 2008; Grimm et al. 2015, 2016), probably because it is sufficient in most cases and is commercially available already linked to dyes such as TMR. However, some published data has also used the Halo(O4) ligand with Dy633 and IR800 (Kosaka et al. 2009).

1.6.3 SNAP-tag and CLIP-tag

HaloTag is not the only covalent coupling technology – other protein systems have been developed in a similar way from different precursor proteins to bind different ligands, for example, SNAP-tag and CLIP-tag. This provides the benefit that they can be used in combination with minimal ligand cross-reactivity, for example with differently attached dyes. However, they are not used in this project and not considered for this work.

1.6.4 Amber codons and click labelling

There are also other methods of targeting fluorophores to macromolecules of interest, for example amber codons and unnatural amino acids (Leinfelder et al. 1990; L. Wang, Xie, and Schultz 2006; J. Wang, Xie, and Schultz 2006; Plass et al. 2012; Schoffelen et al. 2008; Devaraj and Weissleder 2011; Han and Wang 2012), which have been used for SR imaging (Vreja et al. 2015; Sakin et al. 2017; Nikić et al. 2015; Nikić and Lemke 2015) but not SMLM. Nucleic acids have been targeted for SMLM using RNA-FISH (Lubeck and Cai 2012), PAINT (Beliveau et al. 2015) and 5-ethynyl-20-deoxyuridine (EdU) pulse labelling (Zessin, Finan, and Heilemann 2012; Salic and Mitchison 2008). Other methods such as Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR/Cas9) and DNA-FISH are also targeting candidates but have not yet been reported for use with SMLM.

1.7 Strategy to design a new range of flexible STORM probes

This project aims to address some of the disadvantages of STORM methodology by designing and synthesising flexible constructs that may be more suitable to quantitative as well as multicolour imaging.

Flexible multicolour STORM experiments require a system with two modules; a detector module containing an activator and emitter dye in close proximity – to allow imaging; and a targeting module – to target the detector module to a protein of interest, as shown in **Figure 1.15**. STORM can use a variety of activator-emitter pairs in the same sample provided each has a distinct targeting module, thus multicolour STORM imaging of different proteins in a cell is possible, but has only been achieved with Abs (Ricci et al. 2015). To provide flexibility a detector module should therefore be capable of linkage to different activator and emitter dyes.

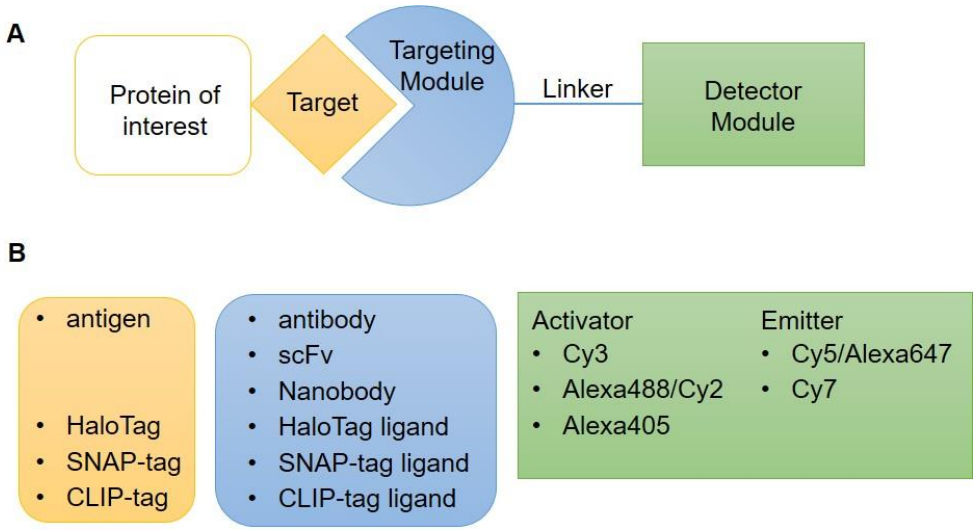


Figure 1.15 – Schematic design of a flexible probe for use in quantitative multicolour STORM imaging. A *The required sections of a probe targeted to a protein of interest.* ***B*** *Possible targeting module and detector module combinations that could be suitable for STORM.*

1.7.1 STORM dyes

Published work on STORM has shown that only a few commercially available dyes are compatible, see **section 1.4.2.4**. For the emitter dye, Cy5 is analogous to Alexa647 and is notably cheaper, so this was selected as the emitter. There are, however, different versions of Cy5 available; the more similar version to Alexa647 is called sulfo-Cy5 because it has sulphonic acid groups to allow water solubility, whereas the other version, Cy5, has no sulphonic acid groups and is less soluble in water. Since cell based experiments are done in aqueous buffers and most of the linking chemistry will be done in water, we chose the sulfo-Cy5 version.

Alexa405, Alexa488 and Cy3 were selected as potential activators for sulfo-Cy5. Cy2 was side-lined in favour of Alexa488 because both are activated by our 488 nm laser, and cannot be used in the same experiment and our lab has had previous experience using Alexa488.

1.7.2 Chemospecific linking of dyes

To create a STORM probe in this manner, it was necessary to target dyes to specific positions to allow a specific one-to-one labelling necessary for potentially quantitative imaging. There are various reactions that are chemoselective – and commonly used for dye coupling to target proteins, although only limited types of reaction are available with purchased dyes.

Commercially available dye reagents for conjugation include the most commonly used *N*-hydroxysuccinimide (NHS) reaction, click chemistry (Kolb, Finn, and Sharpless 2001), maleimide and iodo-alkyl. NHS and click chemistry were selected for use here.

1.7.3 NHS chemistry

NHS groups react with primary amines such as the side chain of a lysine residue as in **Figure 1.16A**, and are often used to label proteins with commercially available NHS-activated fluorophores. This method is widely used with phosphate buffer at pH 8-9 and the target residues are usually lysines, which are commonly available on the solvent-facing surface of most proteins.

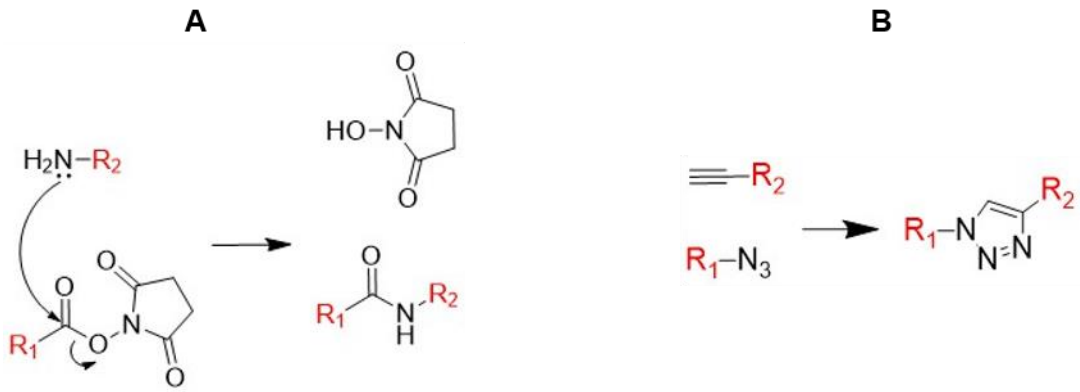


Figure 1.16 – Chemospecific reaction mechanisms showing alkyl groups in red. A *NHS chemistry where R_2 is usually a lysine side chain and R_1 is a fluorophore.* ***B*** *3+2 cycloaddition – "click chemistry" where either R_1 or R_2 can be the target or fluorophore.*

1.7.4 Huisgen dipolar cycloaddition

Many "click chemistry" reactions have been developed, which aim to react small molecules with little to no side product. One of these is the reaction of an alkyne with an azide group in a 3+2 cycloaddition (Kolb, Finn, and Sharpless 2001), see **Figure**

1.16B. The conventional reaction contains the reagents, copper(II) sulphate, ascorbic acid, Tris(benzyltriazolylmethyl)amine (TBTA) and triethylammonium acetate in 55% dimethyl sulfoxide (DMSO). During the reaction, Cu(II) is reduced to Cu(I) by ascorbic acid, and Cu(I) catalyses the 3+2 cycloaddition. To stabilise the Cu(I) ion, oxygen is removed from the buffer by bubbling through nitrogen, and a TBTA ligand is added. Solubilisation of the TBTA organic ligand requires high percentage DMSO, which prevents simultaneous reactions that require aqueous conditions, like NHS. For ease, I shall hence forth refer to the Huisgen dipolar cycloaddition as “click chemistry”.

A more recent development allows copper-free click chemistry via more reactive strained-alkynes (Plass et al. 2011). This reduces copper derived damage *in vivo*, but these are not yet widely available.

1.7.5 Selected chemospecific reactions for a detector module

Alexa405, Alexa488 and Cy3 were only available as NHS ester derivatives, so this targeting chemistry was used. Cy5 on the other hand, was available as an azide as well as an NHS, so this azide was used for click chemistry. The NHS and azide reactions target different reactive groups so should be amenable to specific labelling of residues.

To provide a linker that allows the activator and emitter dyes to sit closely together using NHS and azide chemistry requires a molecule with both amine and alkyne reactive groups. These groups could theoretically be inserted into DNA oligos or Abs as targeting modules, but we are interested in targeting proteins, so DNA oligos primarily used for RNA or DNA-FISH are not viable. Oligos have been linked to Abs for PAINT (Jungmann et al. 2014), but we chose to focus on Abs as targeting modules.

To achieve this specific labelling in a one-to-one manner, we decided to use a new approach involving small synthetic peptides each containing only a single residue for each labelling reaction. This had the advantage that we could include propargylglycine (Pra), which would be more difficult to introduce into recombinantly expressed protein. The peptide was designed to contain one lysine residue, and one Pra residue to allow NHS and click chemistry, respectively. The N-terminus was modified to remove the amino group that might otherwise also react via NHS chemistry.

1.7.6 The targeting module

There are many different methods of targeting proteins of interest for fluorescence imaging as discussed in **section 1.6**. Initially, we investigated several possible targeting modules for possible probes, with the idea of selecting the best later. The "DNA-PAINT" approach relies on washing in and washing out different imager oligos (with attached dyes) to obtain multicolour images and we favour an approach in which the different detectors can be imaged simultaneously, so we chose HaloTag because the HaloPeptides seemed simplest to synthesise, label and purify, and we already had several HaloTag labelled proteins available in different cell strains. We also investigated scFvs because it was hoped that they might be straightforwardly generated from Abs with different known epitopes, and an anti-GFP Nb which had already been well-characterised in other labs (Kubala et al. 2010; Kirchhofer et al. 2010).

scFvs typically need to be expressed in the periplasm of *E. coli* (Martin et al, 2006) because they contain disulphide bonds that cannot form in the reducing environment of the cytoplasm, except in special strains such as SHuffle (Lobstein et al. 2012).(In normal *E. coli* strains, cytoplasmically expressed scFvs may fold incorrectly or aggregate, so functional maturation relies on export to the oxidising environment of the periplasm.) We decided to investigate scFvs against the FLAG and c-Myc tags because if a functional probe was generated it would be useful against a wide range of already tagged proteins

1.7.7 Native Chemical Ligation (NCL) for linking peptides

To use proteins such as scFvs and Nbs as targeting modules for an activator and emitter dye labelled peptide requires another chemospecific reaction to link a peptide to the targeting protein. The chemistry must not have any cross reactivity to the NHS (**section 1.7.3**) and click chemistry (**section 1.7.4**) to allow each reaction to be specific. A suitable method is Native Chemical Ligation (NCL).

NCL is a generalised mechanism between an N-terminal Cys and a thioester group. It was originally developed as a mechanism to synthesise proteins of approximately 100 amino acid residues in length, initially interleukin8 (IL-8; Dawson et al. 1994). The first step of the reaction mechanism (**Figure 1.17**) is a reversible nucleophilic attack by the N-terminal Cys thiol group, which displaces the thioester group, before the N-terminal amine attacks the same group, displacing the sulphur and forming a peptide bond that

is non-reversible at approximately neutral pH. The N-terminal Cys in its reduced state was critical, as disulphide formation prevents the reaction, so thiol reducing agents are usually added. The thioester group also needed to be an effective leaving group, such as an S-benzyl (SBz). Phenylmethanethiol was originally used solely to reduce any disulphide bonds, but was later found to play a crucial role in catalysing the NCL. Because the rate of reaction is relatively slow, benzenethiol was then also found to fulfil this role (Dawson, P E, Churchill, M J, Ghadiri, M R et al. 1997).

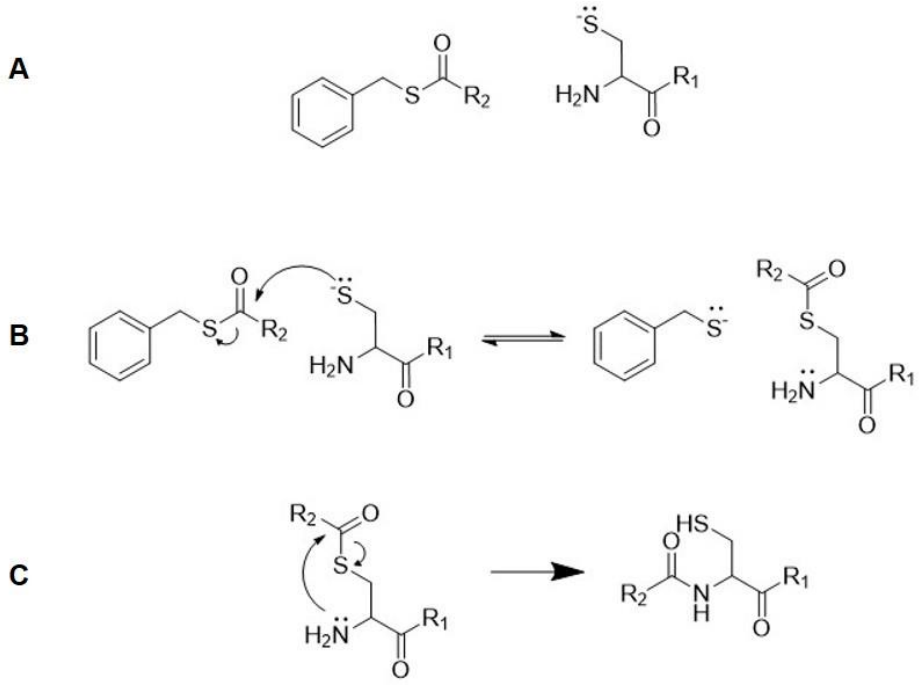


Figure 1.17 – NCL mechanism between peptide R1 with a C-terminal Phenylmethanethiol (SBz) group and peptide R2 containing an N-terminal Cys. A *The starting reagents – an N-terminal Cys and a C-terminal SBz group.* ***B*** *The first reversible step; a nucleophilic attack of the thioester by the sulphydryl group.* ***C*** *The second, irreversible step; an acyl shift produced by nucleophilic attack of the second thioester by the primary amino group of the Cys.*

1.7.8 NCL adaptation for linking STORM dyes to proteins

In the past, NCL and its derivative techniques have been used in peptide synthesis, to make increasing long polypeptide chains, add in synthetic changes like phosphorylated and methylated residues (Bartke et al. 2010) that are otherwise difficult to incorporate when expressing recombinant proteins. In a similar way, NCL could be used here to link Ab proteins to peptides containing activator-emitter STORM pairs.

To facilitate this, we designed a peptide, which could be dual labelled with dyes and then undergo NCL, the structure is shown in **Figure 1.18**. It contained a C-terminal SBz group, similar to those used in the early NCL work (Dawson et al. 1994), and it is referred to as the SBz-peptide. The synthesis scheme for this peptide before NCL requires fluorophore labelling with click and NHS chemistry.

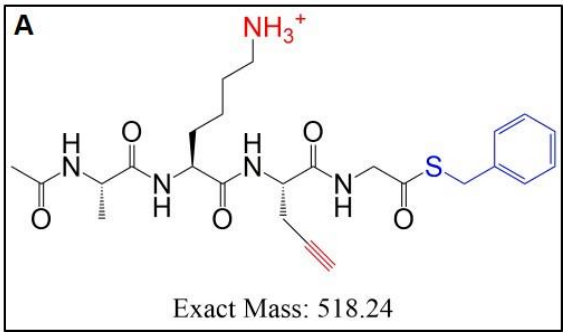


Figure 1.18 – SBz-peptide*. Peptide design capable of linking two different dyes chemospecifically, through the red reactive groups to a protein containing an N-terminal Cys via the blue SBz group in an NCL reaction.*

1.7.9 Periplasmic expression

An important requirement for NCL is the presence of an N-terminal Cys in the protein that the SBz-peptide is to be linked to. In the past this has been done by expressing proteins in the bacterial cytoplasm with modified protease cleavage sequences that, when cleaved, reveal an unprotected N-terminal Cys for NCL (Bartke et al. 2010). However, it is also possible to use the signal peptidases involved in *E. coli* periplasmic export to reveal the N-terminal Cys (Hauser and Ryan 2007). Periplasmic transport is enabled by an N-terminal peptide sequence, such as those from the DsbA or PelB proteins, which target the protein to the respective transport system. The DsbA leader sequence is thought to target the expressed protein as it is translated (Steiner et al. 2006), whereas the sec-dependent PelB leader is thought to act post-translation (Singh et al. 2013; Sletta et al. 2007). After or during transport the signal sequence is cleaved by a membrane bound periplasmic signal peptidase leaving an N-terminus that can be any residue except for those with large, hydrophobic side chains; alanine (Ala) or Cys are most common (K. H. Choo, Tong, and Ranganathan 2007). Therefore, placing a Cys residue immediately after a signal peptide, such as PelB or DsbA, should generate an N-terminal Cys, suitable for linkage to an SBz-peptide.

A further advantage of using periplasmic secretion of the protein is that the periplasm is an oxidising environment, whereas, the cytoplasm is reducing. This means Ab proteins containing disulphide bonds can be expressed and are more likely to fold correctly than with cytoplasmic expression.

1.8 Aims of this project - design of a simple system for multicolour STORM imaging

The overarching aim of this project was to design a flexible one-to-one-to-one STORM system to label target molecules, to do proof of principle experiments to validate the

labelling of target molecules and characterise the photophysical properties required for imaging and then to ask biological questions. To achieve this the main aim has been broken down into four key stages:

i) Make a detector module containing a STORM pair
ii) Make a targeting module and link the two modules
iii) Test the complete system for target binding
iv) Test the system for quantitative experiments in comparison to conventional methods including Abs

If this approach was successful for more than one targeting system and STORM pair it would then be desirable to extend this to other targeting and STORM pairs to achieve quantitative, multicolour imaging in cells.

2 Design and synthesis of a HaloTag compatible STORM probe

2.1 Probe design

As described in **section 1.6.2**, one of the commonly used systems for targeting dyes to proteins of interest is HaloTag. This system is simple because it just requires the HaloTag ligand to be labelled with the desired dyes and added to permeabilised cells expressing a HaloTag fusion. This system was chosen over the similar SNAP and CLIP systems because there were more cell lines readily available with HaloTag fusion proteins and we had previous experience in labelling HaloTag ligands.

This project investigates the possibility of using STORM probes in cells since two dyes (e.g. Cy3 and Cy5) have not previously been linked to a HaloTag ligand. For this reason, both the Halo(O2) and the longer Halo(O4) ligand were used in case the shorter one was unable to dock due to steric hindrance from the dyes.

The main aims of this chapter are:

- To synthesise dual labelled Halo(O2) and Halo(O4) ligands for characterisation
- To characterise the dual labelled ligands for target binding

2.2 Methods

As introduced in **section 1.7.3** and **1.7.4** described in detail in **section 6.6**, there are several chemospecific reactions that have been used here for covalently linking dyes to a targeting module. Our synthesis scheme has used both the NHS and click chemistry reactions for dual labelling.

The click chemistry reaction links a dye with a functional azide group and a molecule with an alkyne group. The reaction contains 50% DMSO, Copper(II) sulphate, TBTA, ascorbic acid and triethylamine acetate. Oxygen was removed by bubbling nitrogen through the solution as described in **section 6.6.1**. DMSO is an organic solvent, and at this high concentration the solution phase is partly organic and not truly aqueous.

The NHS reaction links a primary amine such as the lysine side chain to an NHS activated carboxylic acid, under aqueous conditions with a 200 mM sodium bicarbonate buffering system at pH 7-9, as described in **section 6.6.2**. It is not possible to carry out the NHS and click chemistry reactions sequentially, due primarily to the need of 50% DMSO for click chemistry and aqueous conditions for the NHS reaction. Therefore, reverse phase HPLC (RP-HPLC) was employed to first purify the labelled product of the click reaction before starting the NHS coupling reaction.

RP-HPLC used a C-18 guard column and C-18 column with first buffer of aqueous 0.1% trifluoroacetic acid (TFA) and a gradient of 100% acetonitrile (ACN) with 0.1% TFA. Separation was observed using a single wavelength detector to monitor the dyes' absorbance. Fractions were collected by hand as described in **section 6.6.3**. A vacuum concentrator was used to lyophilise RP-HPLC peaks as described in **section 7.6.3**.

Matrix-assisted laser desorption ionisation mass spectroscopy (MALDI-MS) was used for validation of synthesis and purification of HaloPeptides was performed on site at the Department of Biochemistry facility by Dr Len Packman. All masses are quoted as monoisotopic values for simplicity.

Labelling of fixed Mis6-GFP *S. pombe* cells used 4% (w/v) formaldehyde for pre-labelling and post-labelling fixation. Cell membrane permeabilization was achieved with zymolyase and a Triton-X-100 wash, followed by labelling reagent in the presence of 0.5% BSA, 150 mM NaCl and 0.1% Tween-20, and washes to remove excess labelling reagent, as described in **section 6.8.1**.

Ensemble and SM microscopy of the labelled *S. pombe* cells was carried out on the same instrumental setup as, described in **section 6.8.3**.

2.3 Results

2.3.1 HaloTag ligand design

To use HaloTag as a targeting module for a STORM probe, a peptide with either the Halo(O2) or Halo(O4) ligand at the N-terminus (hereafter termed 'HaloPeptide'), was designed with lysine (Lys) and Pra amino acid residues that could be reacted chemospecifically with two different dyes via NHS and click chemistry, respectively, to produce the desired activator-emitter pairs, as described in **section 1.7**. The design of the two HaloPeptides (Halo(O2)Peptide and Halo(O4)Peptide) are shown in **Figure 2.1**.

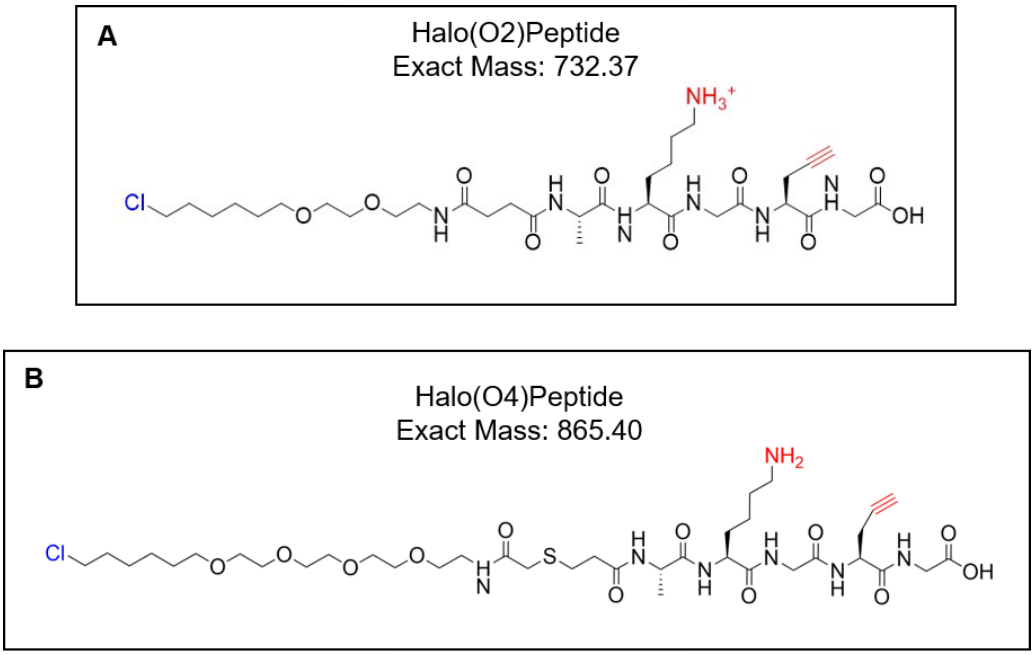


Figure 2.1 – HaloPeptide designs. ***A*** *Halo(O2)Peptide and* ***B*** *Halo(O4)Peptide.*

As discussed in **section 1.4.2.4**, there are only certain dye combinations known to work for STORM, one of which is Alexa488 as the activator and sulfo-Cy5 as the emitter. The azide version of sulfo-Cy5 (sulfo-Cy5-N_3) and the NHS version of Alexa488 (Alexa488-NHS) were purchased and used to label both HaloPeptides using click chemistry and NHS ester chemistry, respectively.

2.3.2 Chemospecific reaction optimisation

It was first important to investigate the two different chemospecific reactions we had chosen to use for linking dyes to the HaloPeptides. The laboratory had no previous

experience with click chemistry and only limited experience with NHS chemistry. It was thought unlikely that both reactions could occur in the same reaction pot, given the differences in manufacturers protocols, so we proceeded with both reactions separately, which required high yield due to only small quantities of the starting HaloPeptide substrates.

2.3.2.1 Click chemistry optimisation

Click chemistry is not commonly used outside of synthetic laboratories and this synthesis requires high yield due to small amounts of starting materials, so optimisation of the reaction conditions was necessary. Various reaction conditions were tested for their effect of yield, including pH, reagent concentration, different types of DMSO and the extent of nitrogen bubbling needed to remove oxygen. We first carried out a pH dependence and found an optimum of 8.7 (**Figure 2.2A).** By varying all these conditions, we optimised the yield considerably as summarised in **Figure 2.2B**, with a final yield of approximately 45%, using a dye concentration of 0.88 mM.

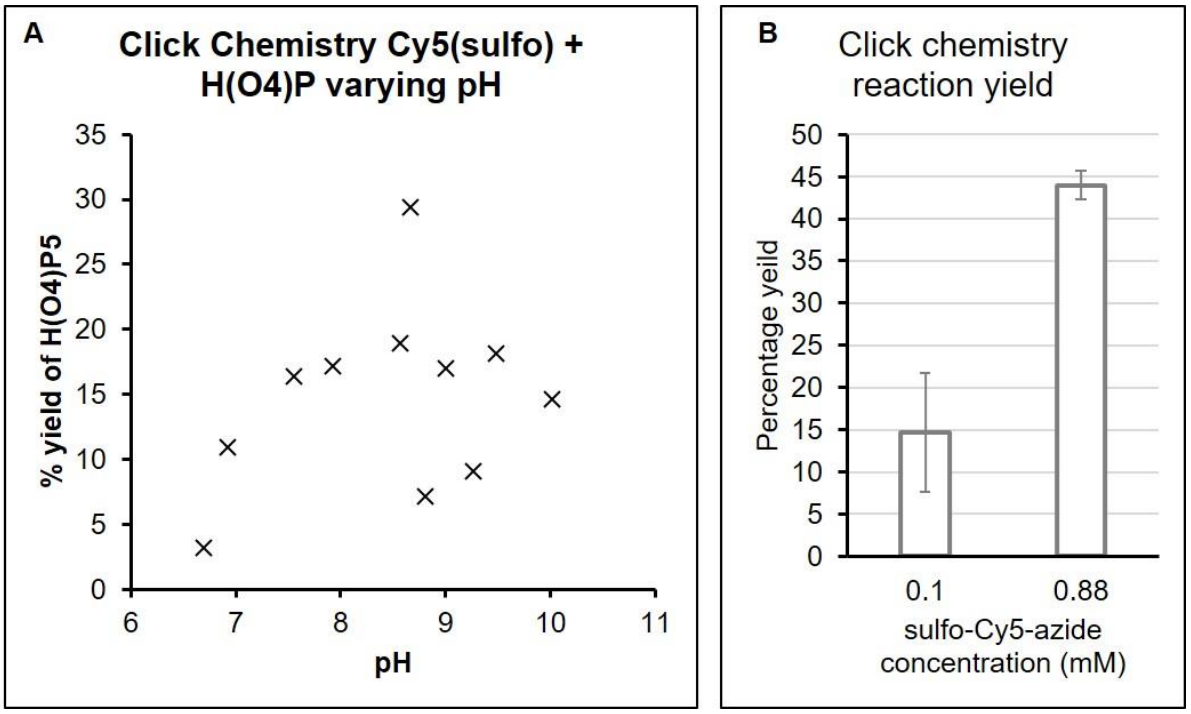


Figure 2.2 - Click chemistry optimisation on 50 µl reaction volumes. ***A*** *pH dependence for the yield of the click chemistry reaction.* ***B*** *Increased yield using higher sulfo-Cy5-N_3 concentration, with anhydrous DMSO and excess dye (as measured as a fraction of Cy5 absorption signal during RP-HPLC).*

2.3.2.2 NHS chemistry optimisation

Whilst NHS reactions are often considered trivial, they are normally carried out on proteins such as Abs in relatively low concentration. We needed high yields so optimisation of the reaction was necessary. NHS esters are sensitive to water hydrolysis, which competes with the desired reaction with the amine. Both reactions

are slower at low pH because the nucleophiles are more protonated, and both the water hydrolysis and the NHS reaction speed up at high pH.

Initial reaction yields were poor – in the range of 2.4-4.5%, this was improved to approximately 40% with pH dependence as shown in **Figure 2.3A**. Yield was better at pH ~7.5, despite the manufacturer's advice of pH 8.5. Optimal pH conditions were achieved by adjusted with 1 M hydrochloric acid (HCl) before NHS-dye addition. The optimal pH also varied depending on the identity of the dye-NHS conjugate, for example, Alexa488-NHS had a pH optimum of approximately 8.3.

The manufacturer of the dyes also played an important role in the stability of the NHS-esters. Considerable difficulty was seen with a period of very low Cy3-NHS reaction yields. This was traced to several batches of the dye, which were mostly hydrolysed upon arrival, despite being lyophilised for delivery. This was proven by RP-HPLC of a brand-new batch as shown in **Figure 2.3B.** The resulting three peaks were subject to MALDI-MS, which identified 2 distinct molecular species. The dominant RP-HPLC peaks 1 and 2 both contained hydrolysed Cy3-NHS ester, whereas, the minor peak 3 was the functional Cy3-NHS. Consultation with the supplier (Lumiprobe) indicated this was a general ongoing problem, so a different manufacturer (GE Healthcare) was sought, and no problems were encountered with their product.

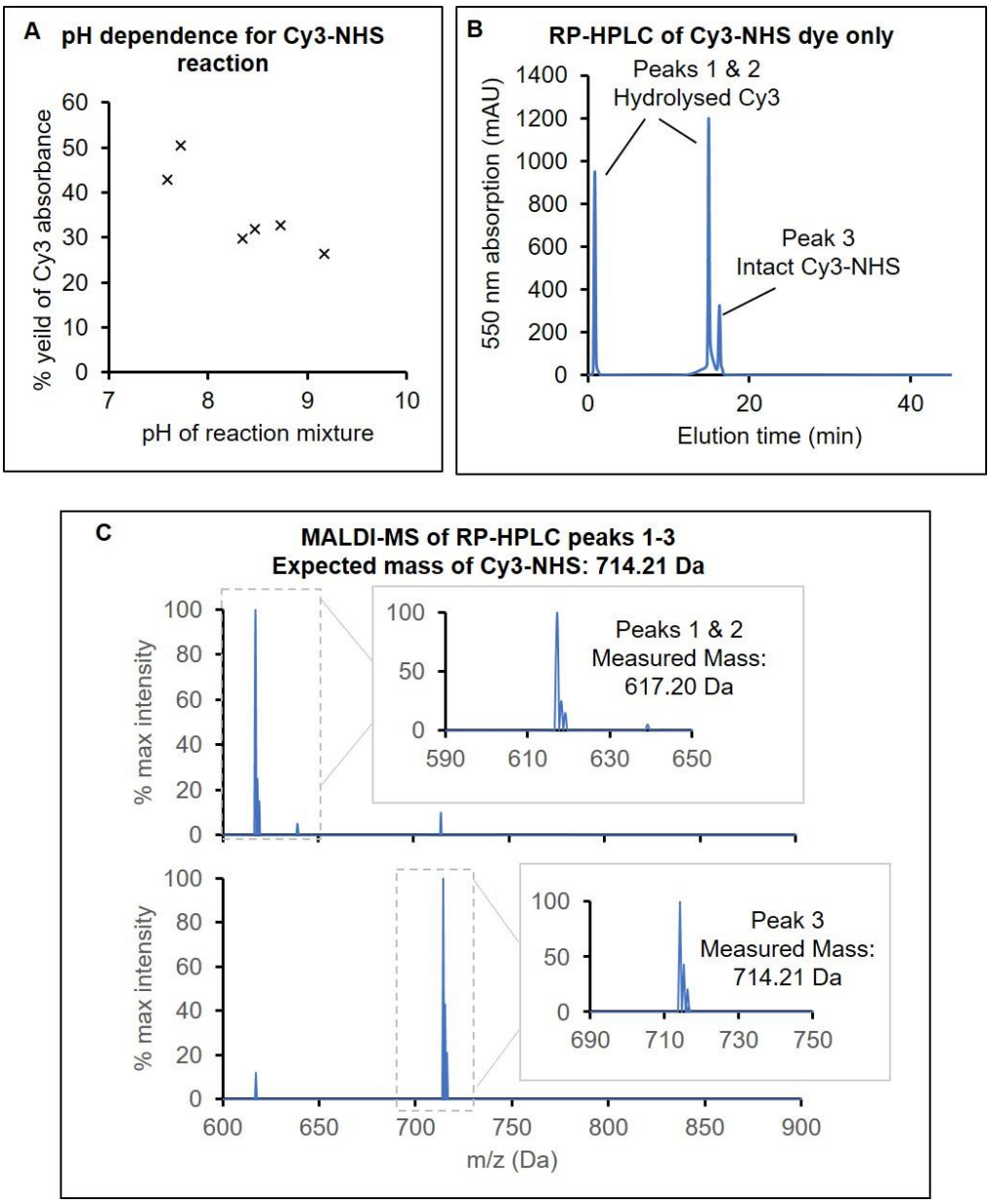


Figure 2.3 – Characterisation of the Cy3-NHS dye. ***A*** *pH dependence for the yield of the NHS reaction.* ***B*** *RP-HPLC trace of brand new Cy3-NHS ester dissolved in DMSO, with the three main peaks labelled.* ***C*** *MALDI-MS spectra for RP-HPLC peaks 1-3 from* ***B****, plotted on the same X-axis. All three have an inset zoomed region on the main set of peaks in that spectrum. Peaks 1 and 2 had a major species of 617.20, Dalton (Da), and Peak 3 had spikes at 714.21 Da.*

2.3.2.3 Order of synthesis

The NHS and click reactions were found to occur under different conditions so they had to be carried out sequentially. Since neither reaction achieved ~40-50% yield, the products of the first reaction must be separated. This was done using RP-HPLC as described in **section 6.6.3**. Although the order in which the reactions were carried out may not have been important, the click reaction was less variable and meant that it was logical to use that first. This also had the advantage that the sulfo-Cy5 was added first, and then this stock of reactant could be coupled to different activators to test and characterise them later.

Since the NHS and click reactions conditions are so different, in order to change the solvent, the products of the click chemistry and RP-HPLC were lyophilised and stored that way until needed.

2.3.3 Synthesising a dual labelled Halo(O2)Peptide for STORM

Firstly, Halo(O2)Peptide was labelled using click chemistry link sulfo-Cy5-N_3. The reagents and desired product are shown in **Figure 2.4A-C**. A control sample containing only the sulfo-Cy5-N_3 was run through RP-HPLC with sulfo-Cy5 containing species detected by an absorbance of 600 nm (the maximum absorbance of Cy5 is ~650 nm) and showed a single peak, indicating a chemically stable dye **Figure 2.4D**. The reaction mixture was then separated in the same way and two separate peaks were produced on the RP-HPLC chromatogram in **Figure 2.4D**, which represent different chemical species produced - Halo(O2)Peptide-sulfo-Cy5 and unreacted sulfo-Cy5-N_3. The desired product was verified by MALDI-MS and had a yield greater than 50%. The two products were lyophilised using a vacuum concentrator.

The lyophilised Halo(O2)Peptide-sulfo-Cy5, was then labelled with Alexa488-NHS, and the reaction products separated in the same way, using 600 nm absorption to detect Cy5 rather than Alexa488, since unreacted and NHS-hydrolysed Alexa488 was in excess to the Cy5 product. The scheme and results for this step are shown in **Figure 2.5**. It yielded two distinct species that absorb at 600 nm (**Figure 2.5C**), each peak due to a different species. Peak 2 eluted from the C-18 column at the same time as the unreacted Halo(O2)Peptide-sulfo-Cy5, so the first peak was probably the desired reaction product. MALDI-MS of the first peak verified this observation by returning a monoisotopic mass of 1974.64 Da, within experimental error of the expected 1974.62 Da. The yield of this second reaction was poor, but was just sufficient material for analysis.

These results demonstrated the successful synthesis of Halo(O2)Peptide-Alexa488-sulfo-Cy5.

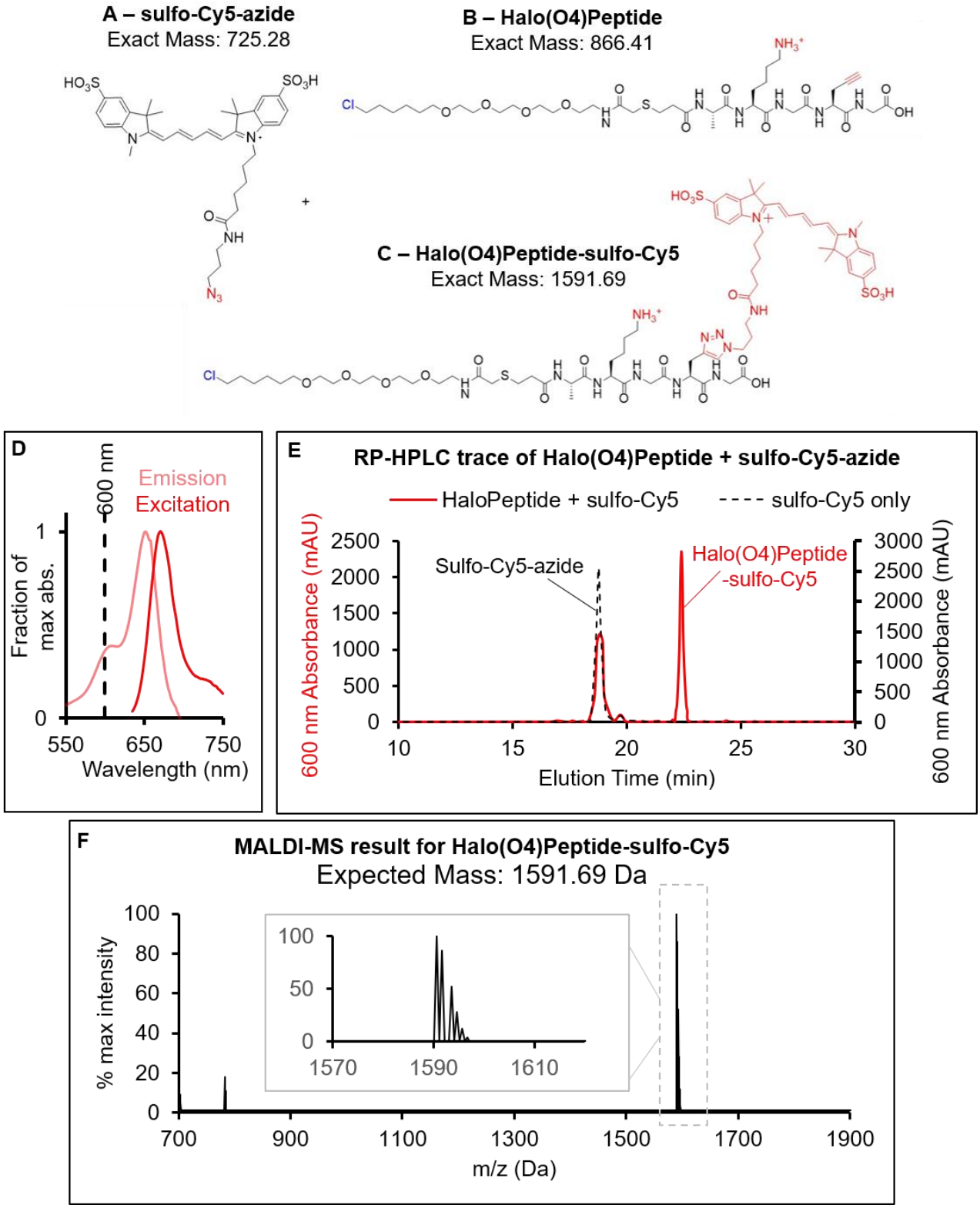


Figure 2.4 – Synthesis of Halo(O2)Peptide-sulfo-Cy5. ***A*** *sulfo-Cy5-N_3* ***B*** *Halo(O2)Peptide,* ***C*** *the desired product Halo(O2)Peptide-culfo-Cy5.* ***D*** *Excitation and emission spectra of Cy5 with the utilized absorbance wavelength, 600 nm, indicated.* ***E*** *RP-HPLC trace of sulfo-Cy5-N_3 in a click reaction mixture with and without Halo(O2)Peptide.* ***F*** *MALDI-MS of the latter reaction peak from A shows the expected mass of 1456.58 Da.*

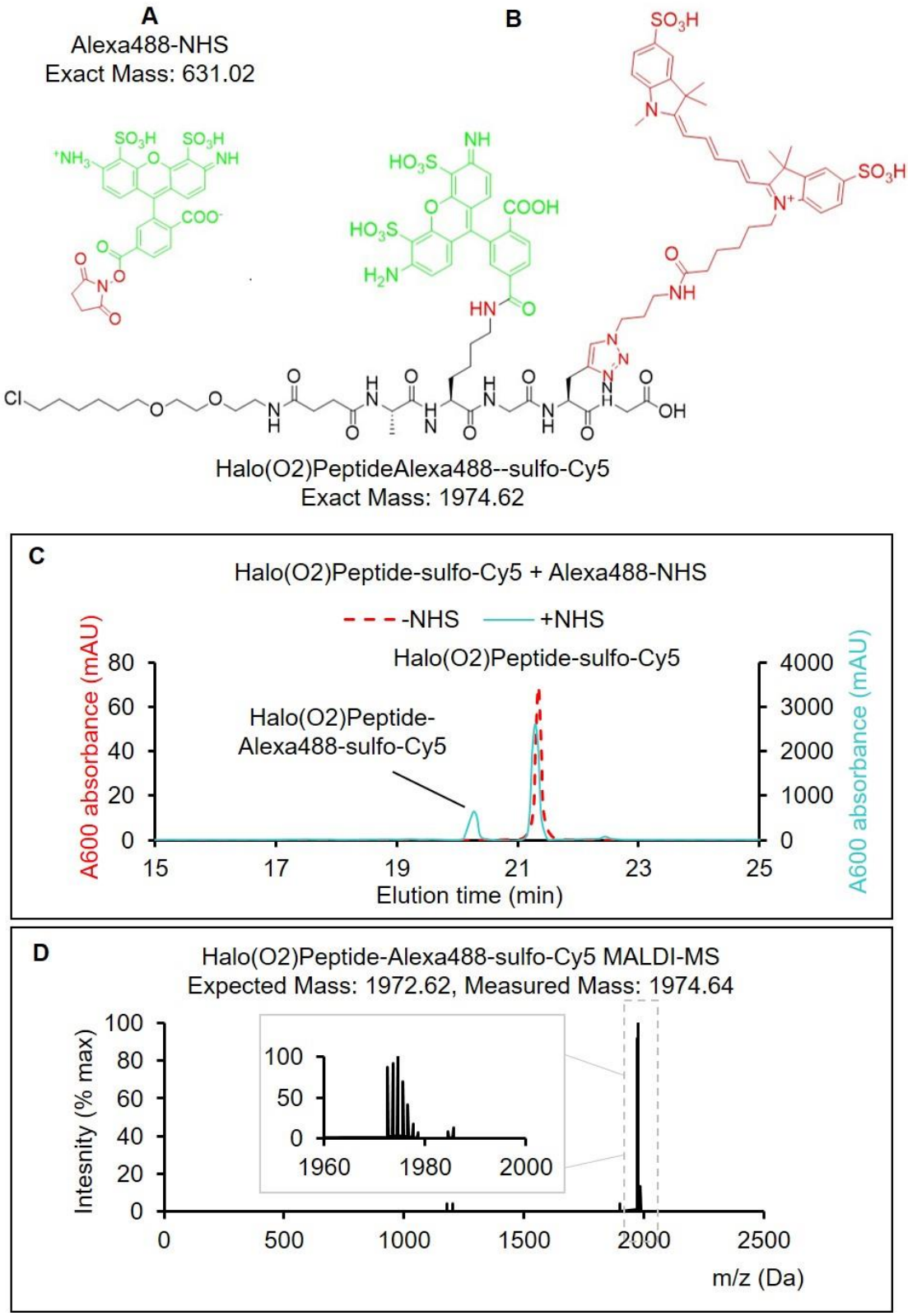


Figure 2.5 – Synthesis of Halo(O2)Peptide-Alexa488-sulfo-Cy5. ***A*** *Alexa488-NHS* ***B*** *The desired product: Halo(O2)Peptide-sulfo-Cy5.* ***C*** *RP-HPLC trace of Halo(O2)Peptide-sulfo-Cy5 with and without Alexa488-NHS.* ***E*** *MALDI-MS of the first reaction peak from* ***B*** *shows the expected mass of 1974.64 Da.*

2.3.3.1 Other synthesised HaloPeptides

Similar reaction schemes were also used to synthesise various other HaloPeptide constructs. These are listed in **Table 2.1**. RP-HPLC traces and MALDI-MS verification

of products can be found in the **Appendix**. To summarise, the click chemistry reaction was robust when using sulfo-Cy5-N_3, and NHS chemistry with Alexa-488 and Cy3-NHS. Alexa405-NHS proved difficult to link to the HaloPeptides partly because it has no visible colour in aqueous solution, which makes following it difficult and RP-HPLC of the dye alone showed that it produced multiple peaks, which were difficult to identify and separate from reaction product peaks. Since two other activators were successful 405-NHS was not pursued further.

Synthesis using Cy5-N_3 (no sulphonic acid groups) was also successful, but the hydrophobic nature of this dye made working with it difficult. It produced broad peaks in RP-HPLC and was not fully soluble in the NHS reaction buffer, so it was not pursued.

Dyes added	**Halo(O2)Peptide**	**Halo(O4)Peptide**
Sulfo-Cy5	**Synthesised**	**Synthesised**
Alexa488 Sulfo-Cy5	**Synthesised**	**Synthesised**
Cy3 Sulfo-Cy5	**Synthesised**	**Synthesis failed**
Alexa405 Sulfo-Cy5	**Synthesis failed**	**Not Attempted**
Cy5	**Not Attempted**	**Synthesised**

Table 2.1 – Attempted synthesis of two different HaloPeptides with sulfo-Cy5 and various activator dyes.

Control HaloTag ligands were also labelled using Atto655-NHS and Alexa647-NHS. The labelled products were also purified by RP-HPLC and verified with MALDI-MS. These were used as positive controls for the labelling HaloTag in a biological context because Atto655 has been published as being used in this way, so must bind HaloTag effectively in cells (Wilmes et al. 2012), and Alexa647 was also expected to work.

2.3.4 HaloPeptide binding validation

To investigate the binding efficacy of the HaloTag protein to its dual labelled ligands; Halo(O2)Peptide-Alexa488-sulfo-Cy5 and Halo(O4)Peptide-Alexa488-sulfo-Cy5, were tested in two different cellular systems – an *S. pombe* strain containing Mis6-HaloTag

(Mis6-Halo) and haploid mES cells with CENP-A labelled with Eos-HaloTag (CENP-A-Eos-Halo). Both Mis6 and CENP-A are centromeric proteins, which means almost all the of these proteins in the cell are localised to the centromeres, which is useful for testing binding.

2.3.4.1 *S. pombe* centromeric proteins Mis6 and Cnp1 as test systems

S. pombe is a useful model organism for testing protein localisation because it has 3 chromosomes, the centromeres of which assemble with many associated inner kinetochore proteins including Mis6 and centromeric histone H3 variant Cnp1 (CENP-A in mammals). This means that the majority of Mis6 and Cnp1 is localised to a small region in the centre of a G1 phase cell (or two regions in the case of a G2 cell because the DNA segregates before mitosis in *S. pombe*; Lando et al. 2012).

When Mis6 or Cnp1 is fluorescently labelled, for example with GFP or with a dye via HaloTag, a single bright focus of light approximately 200 nm across is observed under a wide field fluorescent microscope. S*. pombe* strains Mis6-Halo, GFP-Cnp1, and PA-mCherry-Cnp1, were used as test systems for HaloTag ligands, anti-GFP Nb derivatives and anti-PA-mCherry Nb derivatives. These are good test systems because the fluorescence of the STORM pairs should be very specifically localised to the centromere, which is easily distinguishable. Any non-specific binding should be evident.

2.3.4.2 Haploid mES cell CENP-A-Eos-Halo

Haploid mouse cells also have condensed centromeric regions with many proteins like CENP-A associated there and nowhere else in the genome. However, they have 20 chromosomes in a haploid cell, and they do not cluster together, like in *S. pombe*. Nevertheless, when fluorescently labelled, they are clearly visible as bright foci in the nucleus. Imaging the green form of Eos combined with the fluorophores attached to the HaloPeptides, should be seen in the same foci, if covalent labelling of HaloTag is achieved. Successful labelling would also indicate if these constructs might be useful in mammalian cells.

2.3.4.3 Halo(O2)Peptide-sulfo-Cy5 labelled *S. pombe* Mis6-Halo

We reasoned that a good positive control for HaloPeptide binding was to use the singly labelled version first. This is structurally more like HaloTag ligand labelled with a single dye that has been shown to work previously (Wilmes et al. 2012; Grimm et al. 2015, 2016). The results of the labelling are shown in **Figure 2.6A**. In most of the cells one

or two strong foci were seen per cell when imaging the Cy5 channel. When a profile a long a line was used to measure Cy5 fluorescence a clear peak was observed for the focus, with a maximum signal-to-noise ratio of approximately 4. This is what is expected for good labelling, indicating that docking and the covalent linking are not a problem for this construct.

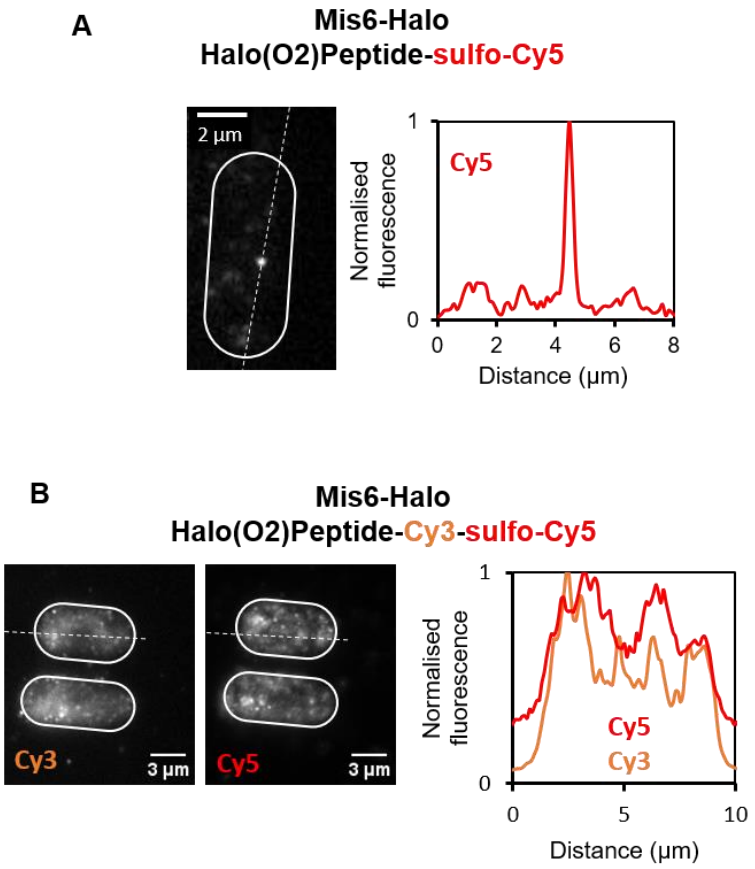


Figure 2.6 – Labelling of S. pombe Mis6-Halo with Halo(O2)Peptide-sulfo-Cy5 and Halo(O2)Peptide-Cy3-sulfo-Cy5. A *Cy5 image of an S. pombe cell stained with Halo(O2)Peptide-sulfo-Cy5 in G2 phase, with a single centromeric focus in the middle and a profile of fluorescence shown through that focus with the dashed line.* ***B*** *Staining of two Mis6-Halo S. pombe cells with Halo(O2)Peptide-Cy3-sulfo-Cy5. Two panels show fluorescence detected in the Cy3 and Cy5 channels, and a fluorescence profile for both channels was taken along the dashed line indicated in the top cell. Scale bars are shown at either 2 or 3 µm, Cy3 fluorescence was detected with a 561 nm laser, and band pass filter, whereas, Cy5 was detected with 641 nm light and a 640 long-pass filter.*

2.3.4.4 Halo(O2)Peptide-Cy3-sulfo-Cy5 cannot label *S. pombe* Mis6-Halo

Halo(O2)Peptide-Cy3-sulfo-Cy5 was also used to label Mis6-Halo cells. Some representative cells are shown in **Figure 2.6B** imaged in bulk in the Cy3 and Cy5 channels to attempt to detect both the activator and emitter. Both bulk fluorescence images qualitatively show a general spread of fluorescence throughout the cells, with some local increases in intensity seemingly randomly dispersed, unlike in **Figure 2.6A**. A profile was taken through one of the cells (**Figure 2.6B**), which showed more quantitively that there was no single clear fluorescent peak, and that background

fluorescence was high along that line in both channels. This high background, above the extracellular background indicated non-specific binding of the Halo(O2)Peptide throughout the nucleus and cytoplasm.

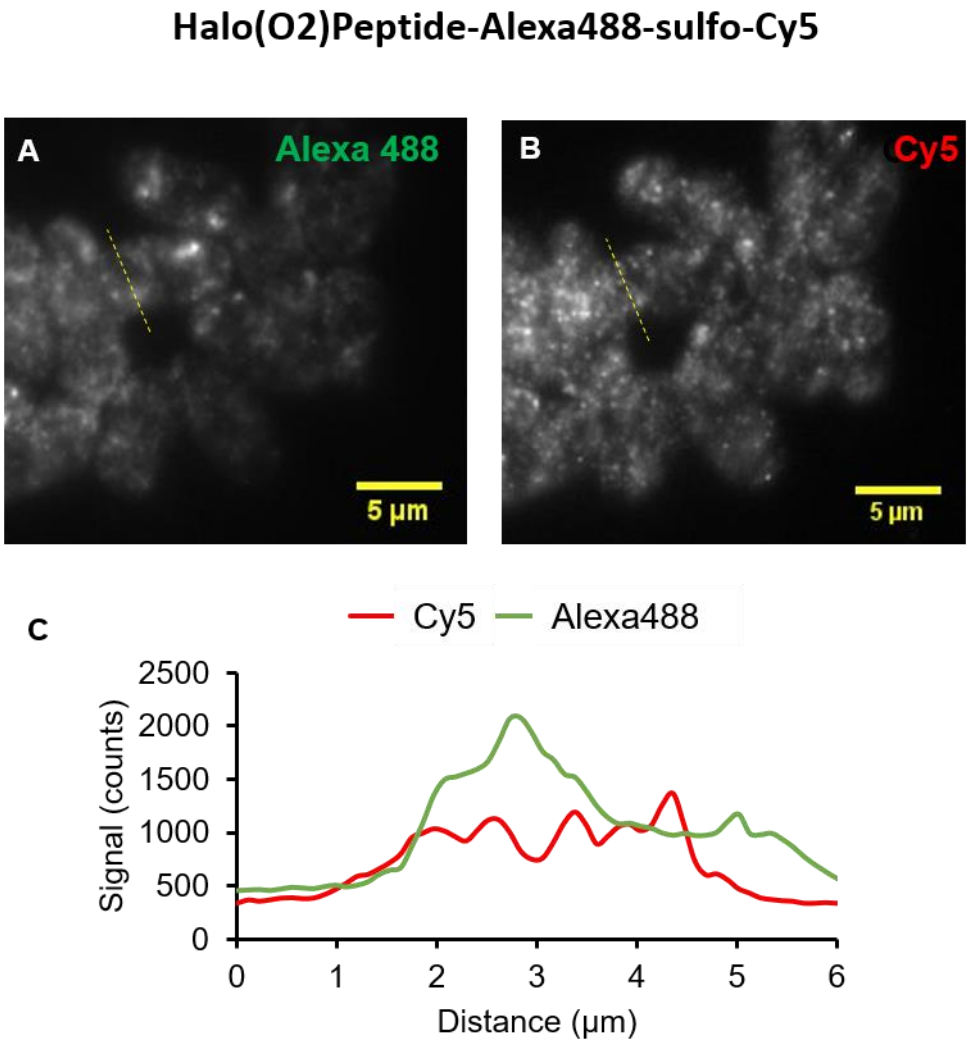


Figure 2.7 – Labelling of mES cells containing CENP-A-Halo with Halo(O2)Peptide-Alexa488-sulfoCy5. A *Imaged in the 488 nm channel with a band pass filter* ***B*** *Imaged with the 641 laser and a 641 long pass filter.* ***C*** *An intensity profile across the same cell for Alexa488 and Cy5 separately.*

2.3.4.5 Halo(O2)Peptide-Alexa488-sulfo-Cy5 could not label mES cell CENP-A-Eos-Halo

The other dual labelled Halo(O2)Peptide containing Alexa488 instead of Cy3 was tested for HaloTag binding in mES cells instead of yeast. A representative colony is shown in **Figure 2.7A and B**. The cells showed no evidence of strong fluorescent foci as would be expected for specific labelling of CENP-A-Eos Halo, and the cross section in **Figure 2.7C** shows both Alexa488 and Cy5 fluorescence spread through the cells, but not present outside them, with no intense peak quantitively evident. This indicates that Halo(O2)Peptide-Alexa488-sulfo-Cy5 is unable to bind to the HaloTag protein in

this construct, despite the Halo(O2)Peptide-sulfo-Cy5 performing well. This may be due to non-specific binding to other proteins in the cells, possibly via charge interactions as discussed later in **Section 2.5.**

2.3.4.6 Halo(O4)Peptide binding assays

The Halo(O2)Peptides seemed unable to bind HaloTag when dual labelled. This was thought to be due to the short linker length, so constructs with the longer Halo(O4)Peptide were also tested for specific linking to Mis6-Halo. The longer linker, was hypothesised to produce less sterically hindrance, and so provide a better chance of specific binding.

A more all-inclusive set of positive controls was also used: Halo(O4)ligand-Atto655, which has been published for in cell imaging (Wilmes et al. 2012) and Alexa647 (Halo(O4)ligand-Alexa647).

2.3.4.7 HaloTag(O4) ligand-Atto655

Labelling Mis6-Halo with HaloTag(O4) ligand-Atto655 produced bright clear foci in nearly every cell. A representative region of interest is shown in **Figure 2.8**. Little noise was observed from inside most of the cell, indicating a very specific interaction between the HaloTag(O4) ligand and the HaloTag protein, even in the dense context of the centromere.

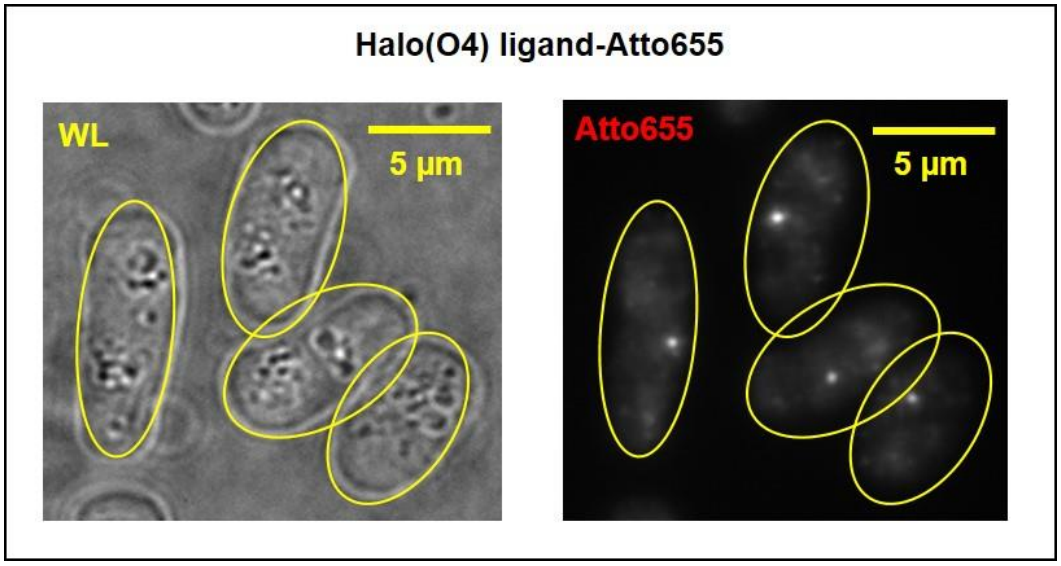


Figure 2.8 – Labelling of S. pombe Mis6-Halo with HaloTag(O4) ligand linked to Atto655. *Widefield image of a several strained cells in WL and Atto655 fluorescence with 641 nm laser.*

2.3.4.8 Halo(O4)ligand-Alexa647

When Mis6-Halo *S. pombe* were labelled with the HaloTag(O4) ligand-Alexa647, there were also clear and bright fluorescent Cy5 foci evident in many cells, an example of which is shown in **Figure 2.9**. Panel **B** shows a dominant bright focus in the middle of the cell as expected for a G1 phase cell. However, compared to Atto655 labelling, there is much more Cy5 fluorescence evident throughout the rest of the cell, which appears to cluster locally in some areas. An intensity trace was taken through this cell to quantitatively demonstrate the non-homogenous fluorescence. At ~8 micrometers (µm) along the line is the clear focus with fluorescence about 3.5-fold above background, but there are other local increases, for example at ~6.5 µm, which indicate that the construct is also binding non-specifically to other areas in the cell and not being washed out during labelling.

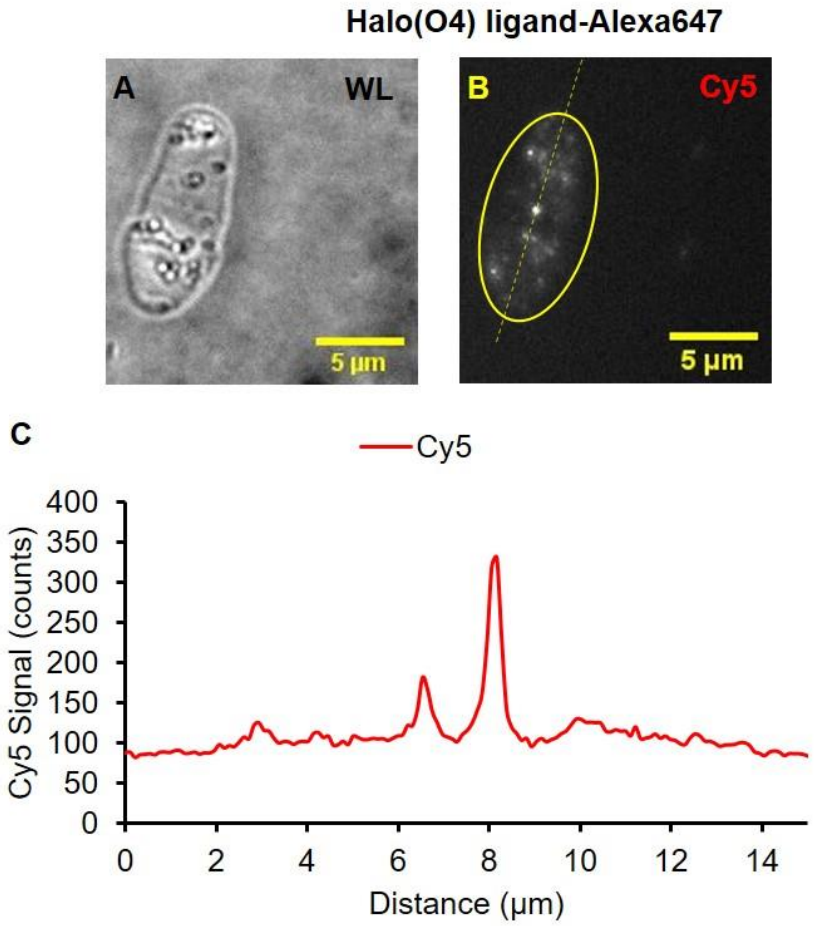


Figure 2.9 – Labelling of S. pombe Mis6-Halo with HaloTag(O4) ligand linked via amine to Alexa647. *Widefield image of a single cell in **A** WL and **B** Cy5 fluorescence with 641 nm laser. **C** Plot of Cy5 fluorescence along the dashed line indicated in B*

2.3.4.9 Halo(O4)Peptide-Alexa488-sulfo-Cy5

Mis6-Halo *S. pombe* cells were also labelled with dual labelled Halo(O4)Peptide containing Alexa488 and sulfo-Cy5 and imaged in bulk fluorescence in the Alexa488 and Cy5 channels to detect both the activator and emitter. An example field of view is shown in **Figure 2.10**.

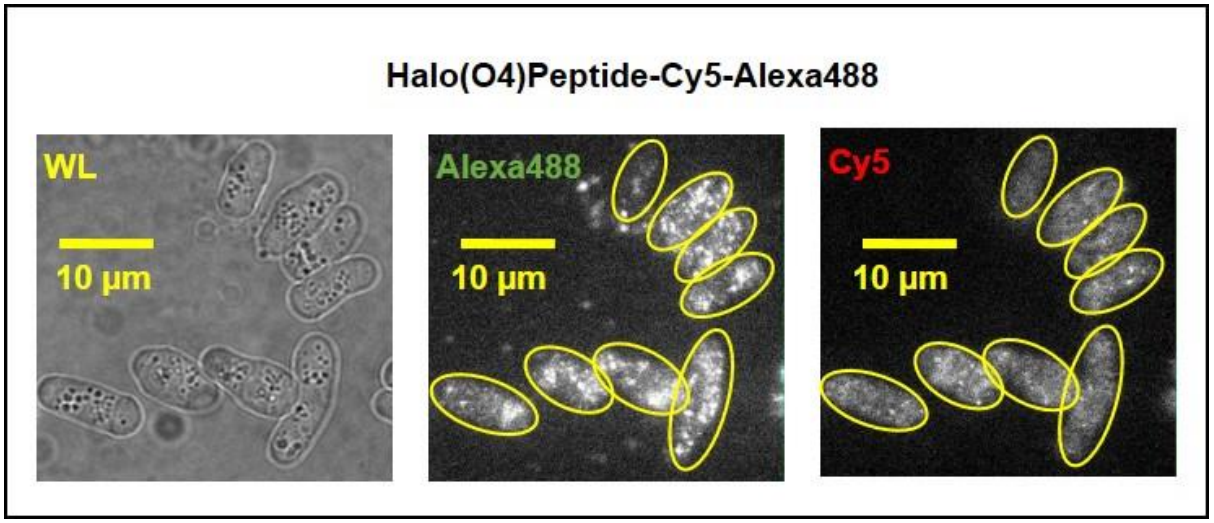


Figure 2.10 – Labelling of S. pombe Mis6-Halo with Halo(O4)Peptides linked to Alexa488 and sulfo-Cy5. *Widefield image of several labelled cells, imaged with **A** WL as well as the **B** Alexa488 and **C** Cy5 channels.*

Both channels show a significant amount of fluorescence within the cells compared to outside the cells, which appears spread across most of the cell in each case. There is no evidence of bright foci distinctive of the centromere. These images therefore, suggest that the dual labelled Halo(O4)Peptide is unable to specifically bind Mis6-Halo, but appears to bind non-specifically throughout the cell, and as such is resistant to washes carried out in the IF protocol.

2.4 Summary

Three dual labelled HaloPeptides were synthesised using click and NHS chemistry to link sulfo-Cy5 and Alexa488 (or Cy3) to either Halo(O2)Peptide of Halo(O4)Peptide. Each reaction was optimised and carried out before RP-HPLC separation and lyophilisation before they were verified by MALDI-MS.

For the first step in characterisation they were all tested for binding efficacy using the *S. pombe* strain Mis6-Halo or the mES cell line CENP-A-Eos-Halo, a summary of the Halo binding experiments for each construct is shown in **Table 2.2**.

.

Construct	**Result labelling Mis6-Halo or CENP-A-Eos-Halo**
Halo(O2)Peptide-sulfo-Cy5	**Successful**
Halo(O2)Peptide-Cy3-sulfo-Cy5	**Failed**
Halo(O2)Peptide-Alexa488-sulfo-Cy5	**Failed**
HaloTag(O4) ligand-Atto655 (positive control)	**Successful**
HaloTag(O4) ligand-Alexa647	**Moderately successful**
Halo(O4)Peptide-Alexa488-sulfo-Cy5	**Failed**

Table 2.2 – Summary of HaloTag binding results with different HaloPeptide and HaloTag ligand constructs.

All dual labelled HaloPeptides failed to produce the expected bright fluorescent foci seen for these centromeric proteins, whereas control labelling with HaloTag(O4) ligand-Atto655 showed strong fluorescent foci in the Atto655 channel, as did singly labelled Halo(O2)Peptide-sulfo-Cy5. Despite rigorous washing in high salt buffer, it seems that the dual labelled HaloPeptides bind non-specifically throughout the cell. Thus, they were classified as unable to bind their target protein in cells.

This result meant that these dual labelled HaloPeptides as a targeting domain for STORM probes was not a viable approach, so other means of targeting were required.

2.5 Discussion

The aim of this chapter was to synthesise dual labelled HaloTag ligand derivatives, potentially capable of STORM, and test their binding efficacy in cells. HaloTag ligands were chosen due to their simplicity; they are easy to label, HaloTag is readily available in useful cell lines and strains, and HaloTag ligands have been shown to bind well in cells. Also some dyes are available commercially already linked to HaloTag ligand and other groups have used them for many approaches like STED (Stagge et al. 2013), *d*STORM (Wilmes et al. 2012; Grimm et al. 2015) and single molecule tracking (Grimm et al. 2016).

Dual labelled STORM HaloPeptides, were unable to specifically label HaloTag fusion of *S. pombe* Mis6 or mESC CENP-A. Whilst the reason for this is not known directly, the control experiments conducted here gave some indication what the cause may be.

In the case of the Halo(O2)Peptides neither dual labelled construct was capable of producing bright foci in the Mis6-Halo or CENP-A-Eos-Halo strains (**Figure 2.6B** and **2.7**). However, the singly-labelled probe has one bright focus in the centre of the cell shown in **Figure 2.6A**, this cell is most likely in G2 or early M phase. This means that two dyes on the Halo(O2)Peptide cannot bind to Halo-Tag, whereas, one dye on the same construct can. There are several potential reasons for this. One is that a second dye sterically hinders the HaloTag(O2) ligand moiety at the N-terminus of the Halo(O2)Peptide, and so prevents binding. If this were the case without any other factors, then the fluorescent construct should be washed out of the cells during the steps that normally remove excess dye, which would result in completely non-fluorescent cells aside from some autofluorescence in the blue-green end of the spectrum. Since **Figure 2.6B** and **Figure 2.7** show considerable fluorescence throughout the cell, this potential docking problem cannot be the only factor.

HaloPeptide fluorescence throughout the cytoplasm and nucleus indicates that there were strong but non-specific interactions between the constructs and something within the cell, which prevent the HaloPeptide linked dyes washing out after incubation. Since the singly labelled construct did not have the same problem, it seems the addition of an extra dye caused or increased those interactions, although the problem was not dye specific as they occurred for both Cy3 and Alexa488. One of the common properties of these dyes is their net negative charge in solution – each has two sulphonic acid groups that are likely to be ionised in solution, as shown in **section**

1.4.2.4. It is possible that these charges cause the non-specific interactions evident in **Figure 2.6B** and **Figure 2.7**, since ionic interactions are strong, and several ionic interactions within the one HaloPeptide may be sufficient to prevent washing out, and prevent free diffusion to the HaloTag target. One possibility is that the labelled HaloPeptides bind histones, which are positively charged at neutral pH, and could explain non-specific nuclear staining, but does not account for non-specific labelling that outside of the nucleus as well, unless there are histones yet to be imported to the nucleus after translation. To investigate this, DNA stains like 4',6-diamidino-2-phenylindole (DAPI) could be used to label the nucleus and investigate whether non-specific binding of the HaloPeptides is confined to the nucleus, in which case histones are the likely cause, or the cytoplasm as well, in which case multiple interactions may be involved.

This speculation is somewhat supported by the Halo(O4)Peptide and HaloTag(O4) ligand evidence, presented in **Figure 2.8-10**. When HaloTag(O4) ligand-Atto655, which is likely to be zwitterionic in solution (**Figure 1.5B**; Wilmes et al. 2012; Grimm et al. 2015, 2016), was used, foci were exemplary (**Figure 2.8**). When the Alexa647, which contains 3 negative charges in solution was used, a clear focus was also evident but there was also evidence of non-specific binding (**Figure 2.9**). The dual labelled Halo(O4)Peptide-Alexa488-sulfo-Cy5, probably containing 4 negative charges in solution, produced high levels of non-specific binding and no evidence of HaloTag interaction (**Figure 2.10**). This trend indicates a correlation between negative charge and non-specific interactions during the IF protocol, which agrees with the Halo(O2)Peptide data.

It seems unlikely that the charges directly prevent HaloTag docking, but it may be that they do cause many non-specific interactions with other proteins that have positively charged surface regions, thus they stick to areas of the cell and do not sample all the space to find HaloTag, especially in the small and densely packed region of the centromere. This may also be a reason why there are few reports of HaloTag linked to sulfo-Cy5 or Alexa647 used for SMLM imaging, other groups may have tried to use for instance HaloTag-Alexa647 and found it unspecific and therefore sought other targeting methods or dyes. It may be of interest to publish this as a technical note, but more data and a detailed explanation would be required. An alternative explanation for the lack of HaloTag-sulfo-Cy5 or -Alexa647 may be because labelled Abs are generally preferable for fixed cell experiments and HaloTag is primarily used in live cell work,

where Alexa647 and sulfo-Cy5 are undesirable because the charge prevents transport across the membrane. Indeed, the live cell tracking and live cell imaging field generally prefers net neutrally charged dyes, hence the development of JF_{549} and JF_{646}, as well as their PA derivatives (Grimm et al. 2016, 2015).

To summarise, the dual labelled HaloPeptides did not bind Mis6-Halo or CENP-A-Eos-Halo specifically, which may be due to steric hindrance, or non-specific interactions of the charged dye groups. Future work is needed to look at the non-specific interactions involved and find ways around them. Without knowledge of why binding is not occurring, future work could focus on different dye combinations with HaloTag ligands that may have less tendency to bind off target. For example, the unsulphonated version of Cy5 and other cyanines may prove more effective at binding HaloTag. As indicated in **Table 2.1**, synthesis of Halo(O4)Peptide with unsulphonated Cy5 was successful, but this construct was very hydrophobic and therefore difficult to react with NHS esters in aqueous conditions. There was insufficient time to fully investigate this possibility.

Whilst labelling of intracellular HaloTag fusion proteins was unsuccessful, it may have been possible to label extra-cellular proteins or regions of membrane proteins with HaloTag. This may have allowed labelling with the dual labelled HaloPeptides without membrane permeabilization and the non-specific interactions observed within the cells. This could have been useful in mammalian cells, which do not have a cell wall like *S. pombe*, and have many membrane proteins involved in signalling. Future work may investigate the use of dual labelled HaloPeptides for STORM on membrane proteins.

Alternatively, similar ligands for different target proteins could be investigated, for example, SNAP-tag and CLIP. Indeed, a STORM probe has been synthesised for SNAP using Cy3 and Cy5 (Dellagiacoma et al. 2010), although this was not extended to multicolour STORM. It may have been the case that other STORM pairs were too difficult to synthesise. It may also have been the case that HaloTag was also tested, but proved not useful because it did not bind and react specifically with the HaloTag enzyme. Again it seems possible that other groups might have observed similar non-specificity and chosen not to publish them.

For this work, an alternative targeting mechanism was required.

3 Design and synthesis of a protein-based STORM probe

HaloTag ligands are not the only option as targeting modules, and since they seem to be unsuitable when linked to 2 dyes, we also investigated Ab based targeting systems. As described in **section 1.6.1**, mammalian IgGs are commonly used for dye targeting in *d*STORM, STORM and PAINT experiments. However, they have several limitations. Their large size may occlude stoichiometric binding; they are bivalent, which may result in under-labelling; and they cannot be labelled in a strictly one-to-one basis, but instead randomly with Poisson distributed statistics around an average number of dyes per molecule.

Therefore, we decided to investigate Ab-derived short-chain variable fragments (scFvs) and nanobodies (Nbs) as potential targeting modules, both of which benefit from small size compared to conventional Abs, are mono-valent, and if expressed and labelled correctly can be made to have exactly one dye per protein. These reagents have been used relatively little for SMLM imaging, and not at all for STORM. The small size is a key advantage because it allows greater penetration of densely packed cellular targets such as heterochromatin, where epitope binding is unlikely to occlude binding of other sites. These advantages mean that a higher proportion of the target protein is likely to be labelled, and it may be possible to do more accurate and precise counting of proteins.

To label an scFv or Nb with a STORM pair of dyes in a precise one-to-one ratio is not possible with random labelling techniques that have been applied in the past. To facilitate this, we designed a peptide like the HaloPeptides (**Chapter 2**), which could be dual labelled with dyes and then ligated using NCL, as described in **section 1.7.7**, to expressed scFvs or Nbs. The peptide structure is shown in **Figure 3.1A**. It contained a C-terminal SBz group, similar to those chosen in the early NCL work (Dawson et al. 1994), and it is referred to as SBz-peptide in this thesis. This scheme also requires the click chemistry reaction, followed by RP-HPLC and lyophilisation, NHS chemistry, RP-HPLC and lyophilisation, to produce a structure like that in **Figure 3.1B**, before the respective proteins can be ligated to by NCL. The Nb and scFv proteins were designed for periplasmic expression, so that signal peptidase cleavage revealed an N-terminal Cys, which is required for NCL.

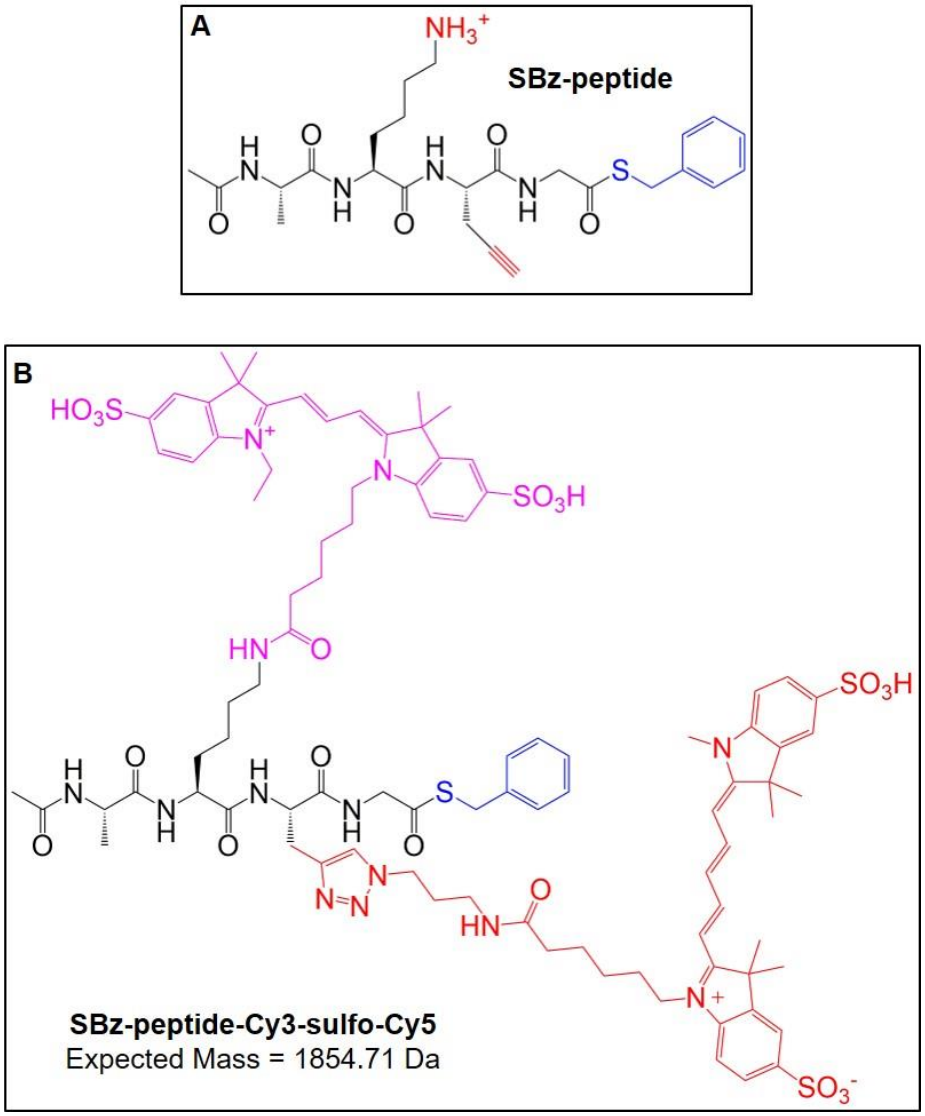


Figure 3.1 – An SBz-peptide detector module for linking to dyes and proteins. A *SBz-peptide structure, with chemospecific reactive groups highlighted in red for dye linkage and in blue for linkage to proteins via NCL* ***B*** *Desired structure of the dye-labelled SBz-peptide labelled with Cy3 (pink) and Cy5 (red) before NCL.*

The key aims of this work were to:

- Synthesise a dual labelled SBz-peptide
- Express and purify one or more scFvs or Nbs, with N-terminal Cys residues
- Use NCL to link the scFv or Nb to the dual labelled SBz-peptide
- Characterise epitope binding in fixed cells.

3.1 Synthesis of a dual labelling SBz-peptide

To produce STORM probes, an activator and an emitter dye must be linked close together. Labelling of the HaloPeptides with sulfo-Cy5 and Alexa488-NHS had been successful, so the same approach was used for labelling the SBz-peptide; first linking the sulfo-Cy5 with click chemistry, and then labelling with Cy3-NHS or Alexa488-NHS. The example purification scheme shown here is for the Cy3-NHS version of the SBz-peptide, which was similar for Alexa488-NHS.

3.1.1 SBz-peptide + sulfo-Cy5

Click chemistry with sulfo-Cy5-N_3 was carried out, followed by RP-HPLC separation of the reaction products. The RP-HPLC trace of sulfo-Cy5-N_3 (**Figure 3.2A**) alone produced a single elution peak at ~14.1 min, as shown in **Figure 3.2C**. After reaction with SBz-peptide, the RP-HPLC trace showed two major peaks plus several smaller ones. The first main elution peak had a similar retention time to the dye only control. The second major peak that eluted at approximately 16 min was thought to be the SBz-peptide-sulfo-Cy5, and this was verified by MALDI-MS, which produced a measured mass of that peak was within error of expected mass of SBz-peptide-sulfo-Cy5 (**Figure 3.2B**), as shown in **Figure 3.2D**. The other smaller peaks were not identified but are likely to be hydrolysis products of the SBz group. The yield of correct product from RP-HPLC was measured by integrating the area under the absorbance peaks, and was found to be 45.55% +/- 2.36% (mean +/- standard deviation of 10 separate purifications), roughly equal to unreacted sulfo-Cy5-N_3 at 44.28% +/- 3.94%, which could be reused subsequently.

These results demonstrated successful synthesis of the required product, with large enough quantity by combining multiple small reactions. They were pooled and lyophilised by vacuum concentrator before proceeding to the next step, because the SBz group was potentially labile in aqueous solution.

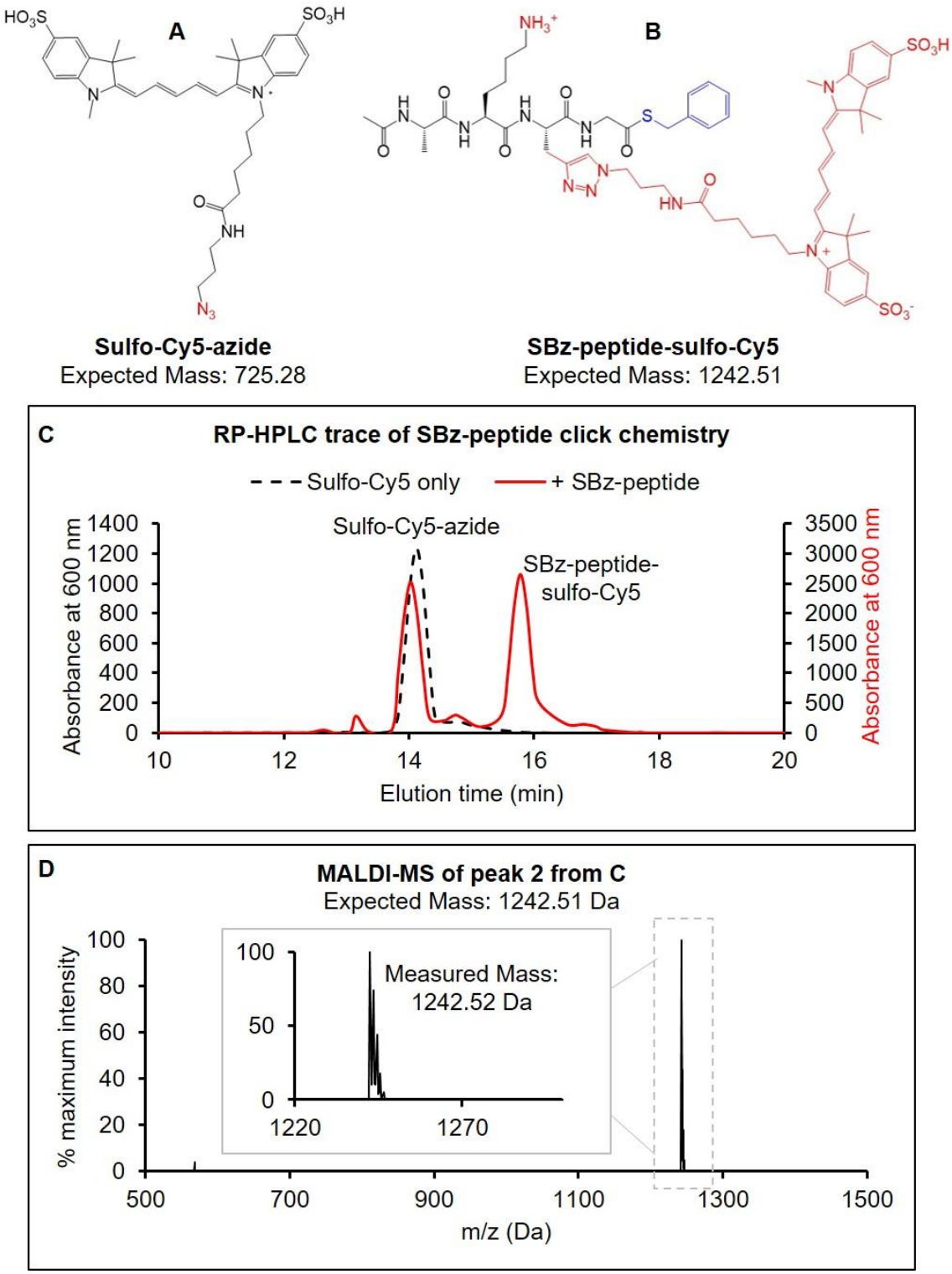


Figure 3.2 – Purification and validation of 2+3 cycloaddition product SBz-peptide-sulfo-Cy5. *Structures of* ***A*** *sulfo-Cy5 reagent and* ***B*** *desired product SBz-peptide-sulfo-Cy5.* ***C*** *RP-HPLC separation of the reaction products compared to the unreacted sulfo-Cy5.* ***B*** *Intact peptide MALDI-MS shows the mass of peak 1 matches to the expected mass of the product SBz-peptide-sulfo-Cy5, with a measured mass of 1242.52 Da, several peaks at +1, 2, 3, 4 Da are also strong.*

3.1.2 SBz-peptide-sulfo-Cy5 + Cy3-NHS

The lyophilised product of the first step, SBz-peptide-sulfo-Cy5, was dissolved in DMSO and reacted with the NHS ester of Cy3 or Alexa488 under aqueous conditions,

as described in **section 6.6.2**. The structure of the dye is shown in **Figure 3.3A** together with the desired product – SBz-peptide-Cy3-sulfo-Cy5 in **Figure 3.3B**. A control containing only SBz-peptide-sulfo-Cy5 and mixture were also separated using RP-HPLC, and example traces are shown in **Figure 3.3C**. The control trace contained two peaks, instead of the one peak expected for this molecular species. MS confirmed this to be due to the hydrolysis of the SBz group from the SBz-peptide-sulfo-Cy5. This molecular species also absorbed at 600 nm but eluted earlier in the acetonitrile gradient because it lacks the hydrophobic benzene ring. This hydrolysis was seemingly unavoidable in alkaline reactions, but since it was only present in relatively small amounts it was not regarded as a large problem.

The reaction of SBz-peptide-sulfo-Cy5 with Cy3-NHS yielded four major products by RP-HPLC, labelled 1-4 in **Figure 3.3C**. These were likely to be the expected reaction product SBz-peptide-sulfo-Cy5-Cy3 and the unreacted SBz-peptide-sulfo-Cy5 together with their requisite SBz-hydrolysis products. To determine which was which, UV-vis absorbance spectra of the peaks were taken and they were analysed by intact peptide MALDI-MS (**Figure 3.3D**). These results showed that peak 1 was OH-peptide-sulfo-Cy5, peak 2 was OH-peptide-Cy3-sulfo-Cy5, peak 3 was unreacted SBz-peptide-sulfo-Cy5 and, the largest, peak 4 was the desired SBz-peptide-Cy3-sulfo-Cy5, with a yield of approaching 50%. This was similar in the case of Alexa488-NHS.

To minimise the amount of SBz-group hydrolysis, NHS reactions were left for as short a time as possible, because a longer reaction seemed to increase the yield of hydrolysis products, and after RP-HPLC the products were placed in the vacuum concentrator as soon as possible, and lyophilised.

In summary, **Figure 3.3** demonstrates the successful synthesis of a dual dye labelled SBz-peptide for linking to proteins. This was also achieved with Alexa488 and Cy5 (not shown).

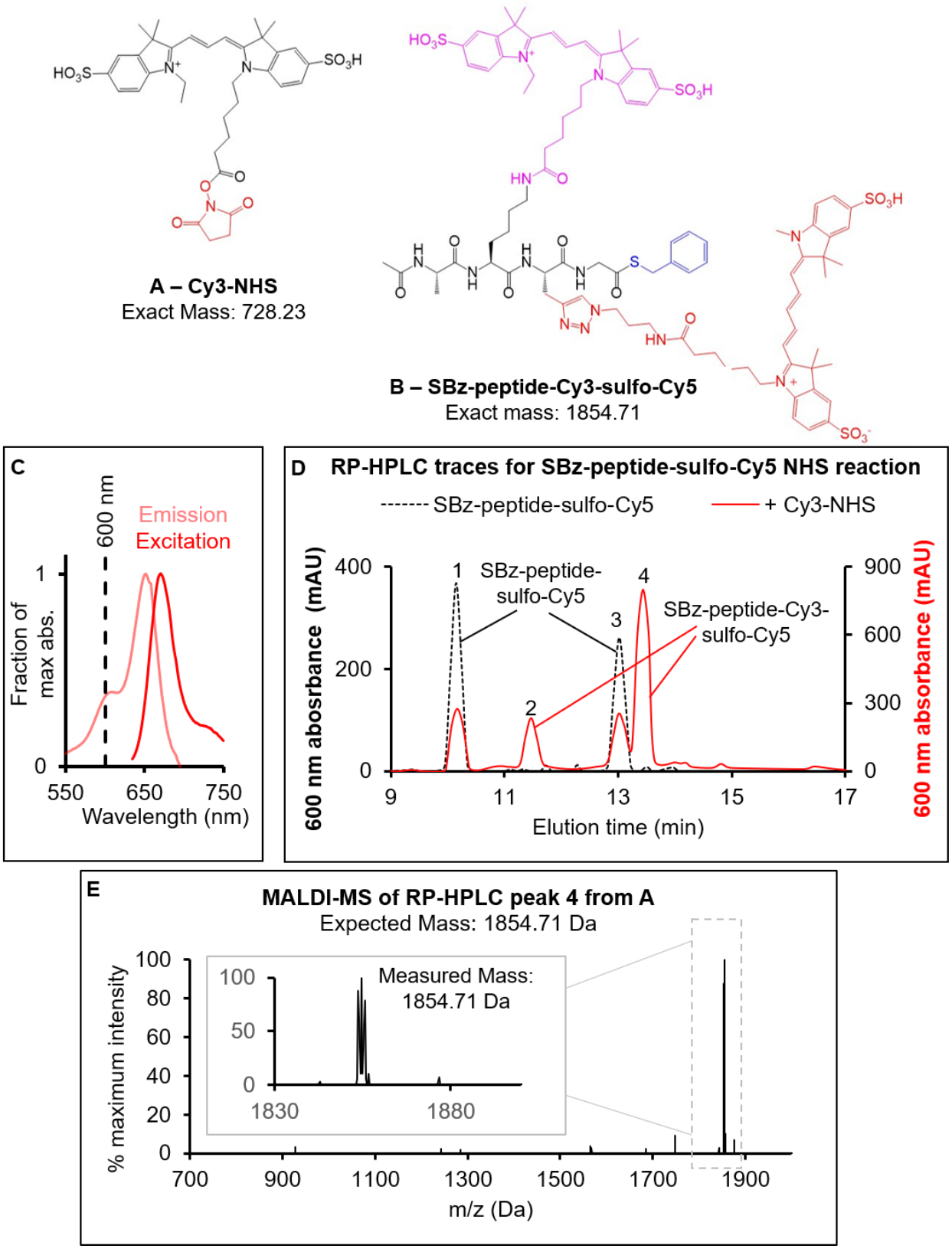


Figure 3.3 – Purifications of SBz-peptide-Cy3-sulfo-Cy5. *Structures of **A** Cy3-NHS reagent and **B** desired product. **C** Excitation and emission spectra of Cy5 with the utilized absorbance wavelength, 600 nm, indicated. **D** RP-HPLC separation of the reaction products compared to the trace for SBz-peptide-Cy5 alone. **E** Intact peptide MALDI-MS of the product from peak 3 has a mass of 1854.71 Da (second isotope peak), the same as that of the desired product from **in B**, with a zoomed scale inset.*

3.2 Targeting protein expression

To make a complete STORM probe using the synthesised SBz-peptide-Cy3-sulfo-Cy5 and SBz-peptide-Alexa488-sulfo-Cy5 and NCL, protein targeting modules were required with an N-terminal Cys. Two scFv and two Nb constructs were designed for this purpose. The aim of this section was to express those constructs in the periplasm and purify them with the N-terminal Cys intact for NCL. Expression was carried out in LEMO21 *E. coli* cells and purification was done using nickel affinity and gel filtration chromatography.

3.2.1 Candidate scFv constructs

Abs are the most widely used tool for labelling cells for STORM and *d*STORM imaging, but they have the drawback of large size, as discussed in **section 1.6.1**. However, scFvs are much smaller and are potentially simpler to produce, as discussed in **section 1.6.1.3**. Therefore, we chose to design two scFvs against c-Myc and FLAG, using published Fab structures (Roosild, Castronovo, and Choe 2006; Krauß et al. 2008).

The two Fab sequences were modified into scFv format using a previously published anti-His-tag scFv scaffold (Kaufmann et al. 2002). **Figure 3.4A** shows the design features of both scFvs, including the crucial N-terminus, linker and C-terminal regions. The designs included 5 and 8 mutations, to the FLAG and c-Myc constructs, respectively, in some solvent facing residues so that V_H residues 11, 16, 61, 62, 63, 64, 106, 182, 183 and V_L residues 9 and 44 (with numbering relative to the original Fab sequences) were all the same as in the anti-His-tag scFv. In some cases, the residues were already the same, hence a different number of mutations were required in each construct. The intention was to increase solubility and stability, based on work with the anti-His-tag scFv (Kaufmann et al. 2002) and an investigation into scFv stability (Miller et al. 2010), which showed that these mutations allow better expression, folding and stability. For example, V_H K106R facilitates a salt bridge that is thermodynamically favourable and helps folding (Kaufmann et al. 2002), and G16E was used in the anti-c-Myc construct because it also increases stability (Miller et al. 2010), but was not used in the anti-FLAG construct as a split strategy. The locations of these mutations relative to the respective Fabs are shown in **Figure 3.4B-D**, and the sequences are aligned by domain in relation to the anti-His-tag scFv and parent Fabs in **Figure 3.5**. An N-terminal signal sequence of either PelB or DsbA was also included as shown in **Figure 3.6A**. They target proteins to the periplasm via the SecYEG complex, as described in **section 1.7.9**; followed by a Cys at the beginning of the V_H domain. The Cys was required for

linkage to the peptide via NCL, and was designed to be present at the N-terminus after signal peptidase cleavage of the leader sequence. A C-terminal hexa-His-tag was also added for purification and immunodetection.

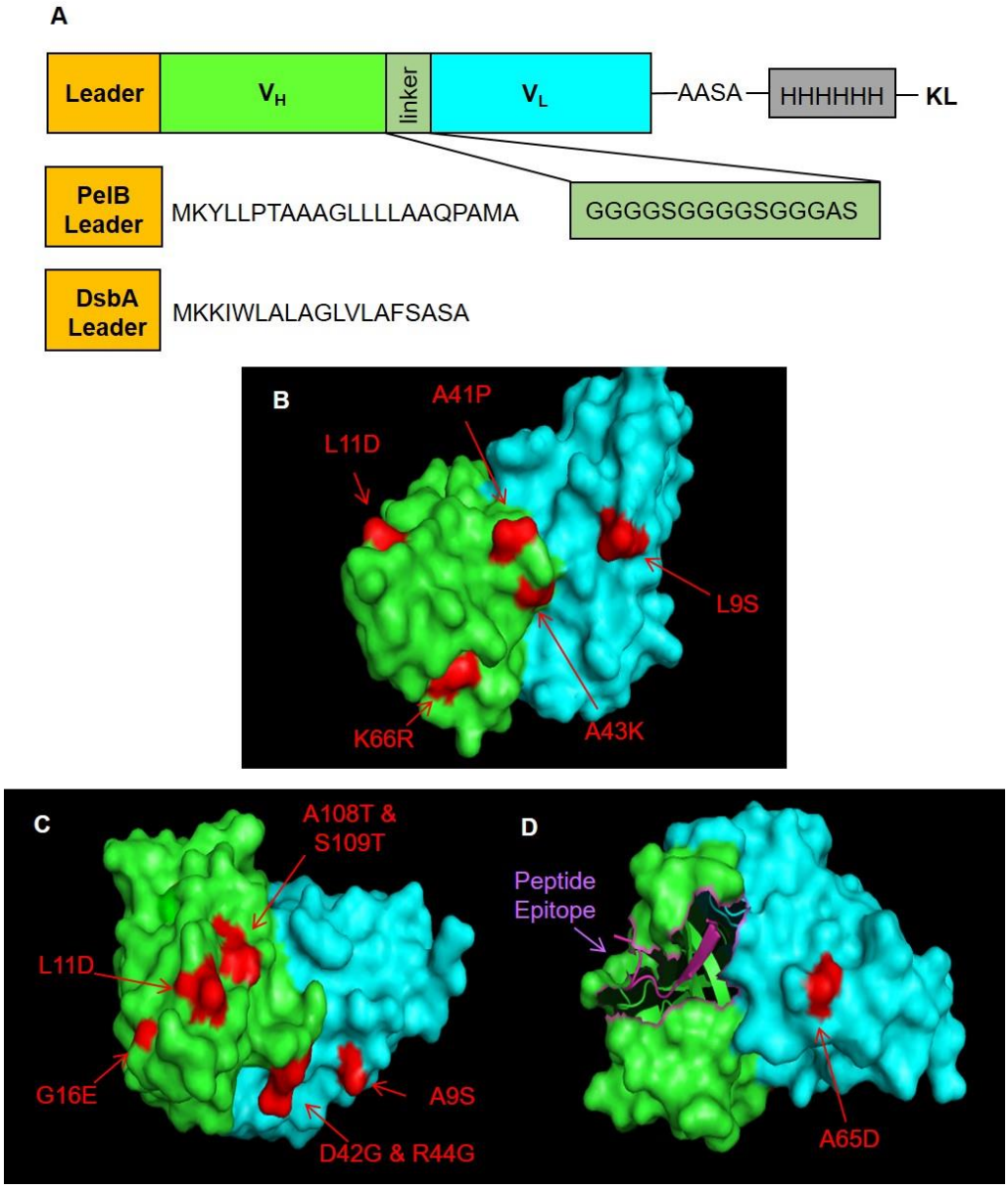


Figure 3.4 – The structure and design of an anti-FLAG and an anti-c-Myc scFv. A *The general arrangement of the domains and sequence features, with the sequences for the two different periplasmic exportation leader sequences.* ***B-D*** *The structures of the initial Fabs, V_H in green and V_L in cyan. The surface rendering is shown, with those residues that have been mutated in the scFv constructs shown in red.* ***B*** *The anti-FLAG Fab structure (PDB: 2G60).* ***C*** *and* ***D*** *Two views of the anti-c-Myc Fab, with the corresponding epitope in magenta (PDB: 2OR9). All of the 8 different mutations are solvent-facing in the structure.*

A - Heavy chain (V_H) sequence alignment

```
FLAG Fab       -EVQLQQSGGELAKPGASVKMSCKSSGYTFTAYAIHWAKQ 39
FLAG scFv      CEVQLQQSGGEDAKPGASVKMSCKSSGYTFTAYAIHWAKQ 40
HisTag scFv    -QVQLQQSGPEDVKPGASVKISCKASGYTFTDYYMNWVKQ 39
MYC Fab        -EVHLVESGGDLVKPGGSLKLSCAASGFTFSHYGMSWVRQ 39
MYC scFv       CEVHLVESGGDDVKPGESLKLSCAASGFTFSHYGMSWVRQ 40

FLAG Fab       AAGAGLEWIGYIAPAAGAAAYNAAFKGKATLAADKSSSTA 79
FLAG scFv      APGKGLEWIGYIAPAAGAAAYNAAFKGRATLAADKSSSTA 80
HisTag scFv    SPGKGLEWIGDINPNNGGTSYNQKFKGRATLTVDKSSSTA 79
Myc Fab        TPDKRLEWVATIGSRGTYTHYPDSVKGRFTISRDNDKNAL 79
Myc scFv       TPGKGLEWVATIGSRGTYTHYPDSVKGRFTISRDNDKNAL 80

FLAG Fab       YMAAAALTSEDSAVYYCARAA----------AAGADYWGQ 109
FLAG scFv      YMAAAALTSEDSAVYYCARAA----------AAGADYWGQ 110
HisTag scFv    YMELRSLTSEDSSVYYCESQS----------GA---YWGQ 106
Myc Fab        YLQMNSLKSEDTAMYYCARRSEFYYYGNTYYYSAMDYWGQ 119
Myc scFv       YLQMNSLKSEDTAMYYCARRSEFYYYGNTYYYSAMDYWGQ 120

FLAG Fab       GTTLTVSS 117
FLAG scFv      GTTLTVSS 118
HisTag scFv    GTTVTVSA 114
Myc Fab        GASVTVSS 127
Myc scFv       GTTVTVSS 128
```

B - Light chain (V_L) sequence alignment

```
FLAG Fab       DVLMTQAPLTLPVSLGDQASISCRSSQAIVHANGNTYLEW     40
FLAG scFv      DVLMTQAPSTLPVSLGDQASISCRSSQAIVHANGNTYLEW     40
HisTag scFv    DILMTGTPSSLPVSLGDQASISCRSSQSIVHSNGNTYLEW     40
Myc Fab        DIVLTQSPASLAVSLGQRATISCRASESVDN-YGFSFMNW     39
Myc scFv       DIVLTQSPSSLAVSLGQRATISCRASESVDN-YGFSFMNW     39

FLAG Fab       YLQKPGQSPALLIYKVANRFSGVPDRFSGSGSGTDFTLKI     80
FLAG scFv      YLQKPGQSPALLIYKVANRFSGVPDRFSGSGSGTDFTLKI     80
HisTag scFv    YLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDFTLKI     80
Myc Fab        FQQKPGQPPKLLIYAISNRGSGVPARFSGSGSGTDFSLNI     59
Myc scFv       FQQKPGQPPKLLIYAISNRGSGVPDRFSGSGSGTDFSLNI     59

FLAG Fab       SRVEAEDLGVYYCFQGAHAPYTFGGGTKLEIKRA 114
FLAG scFv      SRVEAEDLGVYYCFQGAHAPYTFGGGTKLEIKRA 114
HisTag scFv    SRVEAEDLGVYYCFQGSHVPFTFGSGTKLEIKR- 113
Myc Fab        HPVEEDDPAMYFCQQTKEVPWTFGGGTKLEIKRA 113
Myc scFv       HPVEEDDPAMYFCQQTKEVPWTFGGGTKLEIKRA 113
```

Figure 3.5 – Sequence alignment for two candidate scFv with their corresponding parent Fabs and the anti-His-tag-scFv scaffold. A *Alignment of the V_H domain.* ***B*** *Alignment of the V_L domain. Hypervariable loop regions or complementary determining regions (CDRs) are highlighted in yellow and introduced point mutations are in red.*

Both scFv plasmid constructs were purchased with a DsbA signal sequence immediately before the important Cys. An alternative signal sequence in common usage was the PelB leader, so the constructs were both altered using restriction cloning, as described in **section 6.5.1,** to contain the PelB leader so that both signal sequences could be tested. The resulting N-terminal sequences generated by DNA sequencing are shown in **Figure 3.6**.

A

```
DsbA-anti-c-Myc     MKKIWL-ALAG--LVLAFSASACEVHLVESGGDDVKPGES  37
PelB-anti-c-Myc     MKYLLPTAAAGLLLLAAQPAMACEVHLVESGGDDVKPGES  40
                    ** :   * **  *: *  * ******************

DsbA-anti-c-Myc     LKLSCAASGFTFSHYGMSWVRQTPGKGLEWVATIGSRGTY  77
PelB-anti-c-Myc     LKLSCAASGFTFSHYGMSWVRQTPGKGLEWVATIGSRGTY  80
                    ****************************************

DsbA-anti-c-Myc     THYPDSVKGRFTISRDNDKNALYLQMNSLKSEDTAMYYCA  117
PelB-anti-c-Myc     THYPDSVKGRFTISRDNDKNALYLQMNSLKSEDTAMYYCA  120
                    ****************************************

DsbA-anti-c-Myc     RRSEFYYYGNTYYYSAMDYWGQGTTVTVSS             147
PelB-anti-c-Myc     RRSEFYYYGNTYYYSAMDYWGQGTTVTVSS             150
                    ****************************
```

B

```
DsbA-anti-FLAG      MKKIWL-ALAG--LVLAFSASACEVQLQQSGGEDAKPGAS  37
PelB-anti-FLAG      MKYLLPTAAAGLLLLAAQPAMACEVQLQQSGGEDAKPGAS  40
                    ** :   * **  *: *  * ******************

PelB-anti-FLAG      VKMSCKSSGYTFTAYAIHWAKQAPGKGLEWIGYIAPAAGA  77
DsbA-anti-FLAG      VKMSCKSSGYTFTAYAIHWAKQAPGKGLEWIGYIAPAAGA  80
                    ****************************************

DsbA-anti-FLAG      AAYNAAFKGRATLAADKSSSTAYMAAAALTSEDSAVYYCA  117
PelB-anti-FLAG      AAYNAAFKGRATLAADKSSSTAYMAAAALTSEDSAVYYCA  120
                    ****************************************

DsbA-anti-FLAG      RAAAAGADYWGQGTTLTVSS                     137
PelB-anti-FLAG      RAAAAGADYWGQGTTLTVSS                     140
                    ********************
```

Figure 3.6 – Sequence alignment for two candidate scFv N-termini, including the entire V_H domain with two different signal sequences – DsbA and PelB. A *Anti-c-Myc scFv.* ***B*** *Anti-FLAG scFv.*

3.2.1.1 Anti-FLAG and anti-c-Myc scFv expression

Expression and cellular processing of these constructs was studied by Western Blotting using an anti-His-tag Ab, to determine whether they expressed and whether they were successfully secreted into the periplasm. Whilst a well characterised periplasmic scFv [PelB-anti-β-Galactosidase (anti-β-Gal); a gift from John McCafferty], was expressed in both the periplasm and cytoplasm of *E. coli* BL21 cells; the PelB-anti-FLAG, DsbA-anti-FLAG, PelB-anti-c-Myc and DsbA-anti-c-Myc scFvs were only detected in the insoluble fraction of the cell (an example is shown in **Figure 3.7B**). This implies that they aggregated in the cytoplasm, which precluded transport to the periplasm, where folding would occur. This may have been due to insufficient numbers of chaperone proteins that bind the unfolded peptide chains until they were transported to the periplasm, too fast expression, or that the scFv designs are inherently likely to aggregate.

To test the possibility that aggregation was caused by overwhelming of the chaperone and transport systems, we tested multiple conditions that should have reduce the rate of expression and therefore made successful secretion more likely. Methods to do this included lowing the temperature of culture during induction, and using *E. coli* strains like LEMO21, which had a pLEMO21 plasmid with a rhamnose inducible lysozyme gene to inhibit RNA polymerase. Increasing the concentration of rhamnose, therefore, inhibits scFvs expression more.

The same result of scFvs present in the insoluble fraction of the cytoplasm was also obtained under the various conditions designed to slow expression: lower induction temperature (16^0C, 30^0C and 37^0C), rhamnose concentration (0, 0.1, 0.5, 1.0 and 2 M) in LEMO21, and induction mechanism (IPTG or auto-induction). The results in each case were similar to **Figure 3.7B** – the protein was nearly always expressed, but only in the insoluble fraction of the cytoplasm. These experiments suggested that these scFvs aggregated within the cell before the co-translational (DsbA) or post-translational (PelB) chaperones could bind and transfer them to the export machinery. This suggestion was supported by evidence from a large scale DsbA-anti-c-Myc culture that contained the protein (as seen by Western Blot), but which could not be purified on Ni-NTA beads. This indicated that the His-tags were ‘hidden’ from the beads prior to denaturation, but when denatured on an SDS gel, the tags become ‘visible’ to the anti-His-tag Ab.

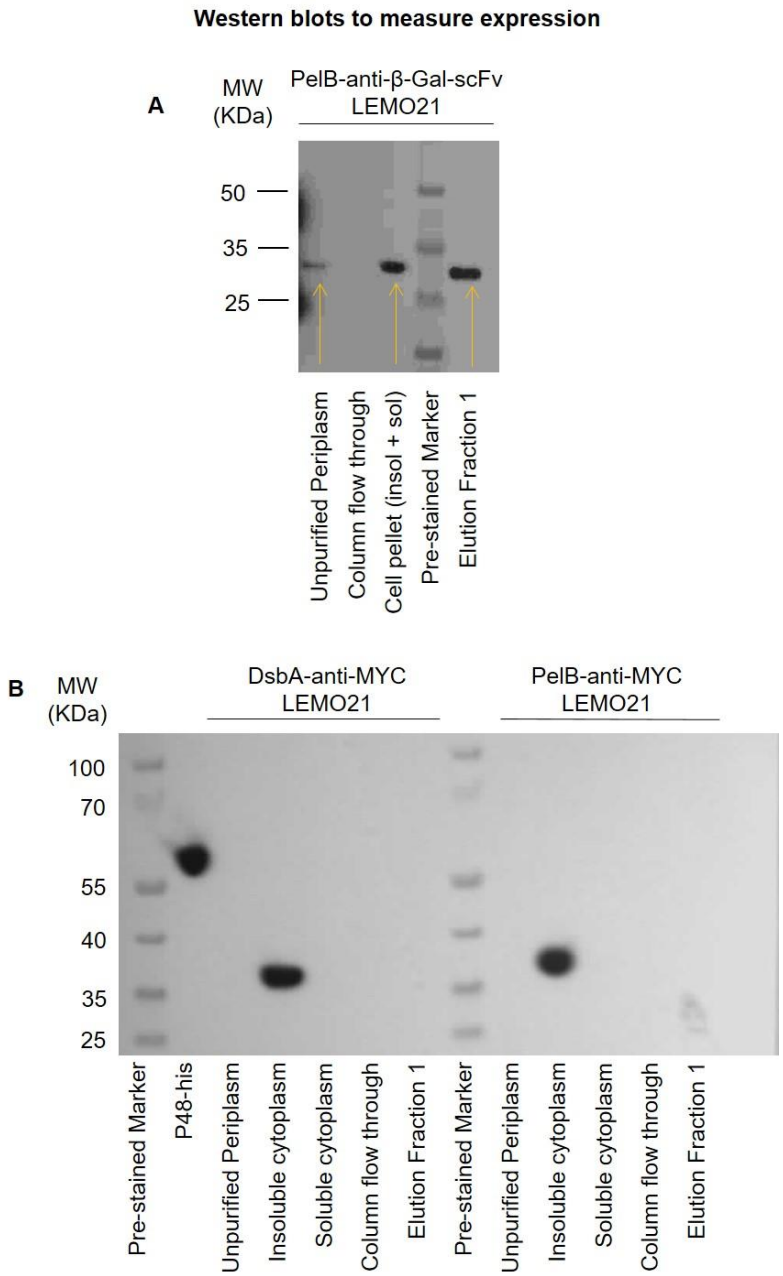


Figure 3.7 – Anti-c-Myc and anti-FLAG-scFvs could not be expressed and exported to the periplasm; they remain in the insoluble fraction. A *Western Blot for positive control protein, PelB-anti-β-Gal-scFv, shows that the method of expression, periplasmic extraction and purification works.* ***B*** *Western Blot of DsbA- and PelB-anti-c-Myc-scFv (~30 KDa) in LEMO21 cells, grown in a 50 mL culture at 16°C overnight; shows that none of the scFv is exported to the periplasm. In both of the examples the positive control for the Western Blot itself is p48 with a C-terminal His-tag.*

To obtain soluble scFvs, Medimmune were approached for help in design and expression of scFvs. They successfully expressed scFvs against FLAG and CHD4 through the conventional isolation of mAbs from serum and affinity maturation, however, by this time other avenues were looking promising, so this was not investigated further.

3.2.2 An anti-GFP Nb

Expression of scFvs as targeting modules for STORM probes proved fruitless, so Nbs were investigate as alternative targeting modules. They were thought to be better prospects because expressed constructs from previous reports were readily available, so there would be no design complications and the single domain nature of Nbs means they are more likely to express and fold correctly than scFvs.

A plasmid for expression of an anti-GFP-Nb (hereafter GFPNb) was obtained directly from the Collins laboratory in the pOPIN-E vector (Kubala et al, 2010), as described in **section 6.5**. This construct had been used for a crystallographic study so was likely to be robustly expressed.

3.2.2.1 Adding N-terminal signal peptides to a GFPNb.

The coding sequence for the original construct was unsuitable for periplasmic expression and subsequent NCL, so it was altered using Gibson Assembly as described in **section 6.5.2**. Both resulting sequences are also shown aligned to the original in **Figure 3.8**, they contain the N-terminal signal sequences PelB or DsbA, and a subsequent CysGlu pair that should be optimal for signal peptidase cleavage.

```
DsbA-GFPNb          MKKIWL-ALAGLVLA--F-SASACEVQLVESGGALVQPGG  36
anti-GFPNb          ----------------------MQVQLVESGGALVQPGG  17
PelB-GFPNb          MKYLLPTAAAGLLLLAAQPAMACEQVQLVESGGALVQPGG  40
                                          :***************

PelB-GFPNb          SLRLSCAASGFPVNRYSMRWYRQAPGKEREWVAGMSSAGD  76
anti-GFPNb          SLRLSCAASGFPVNRYSMRWYRQAPGKEREWVAGMSSAGD  57
DsbA-GFPNb          SLRLSCAASGFPVNRYSMRWYRQAPGKEREWVAGMSSAGD  80
                    ****************************************

DsbA-GFPNb          RSSYEDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYYC  116
anti-GFPNb          RSSYEDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYYC  97
PelB-GFPNb          RSSYEDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYYC  120
                    ****************************************

DsbA-GFPNb          NVNVGFEYWGQGTQVTVSSKHHHHHH  142
anti-GFPNb          NVNVGFEYWGQGTQVTVSSKHHHHHH  123
PelB-GFPNb          NVNVGFEYWGQGTQVTVSSKHHHHHH  146
                    **************************
```

Figure 3.8 – Sequences of the GFPNb with the signal sequences after alignment with the parent gene obtained from the Collins laboratory (Kubala et al. 2010)***.*** ** indicates the same residue and : indicates a chemically similar residue.*

3.2.2.2 Optimisation of Nb expression conditions

GFPNb expression was initially carried out using LB medium and periplasmic extraction and purification. It was found that both the PelB or DsbA signal sequence

facilitated successful export into the periplasm and that they were secreted or leaked into the extracellular medium during cell culture. It was simpler and cheaper to purify the protein from the medium than to do a conventional periplasmic extraction. All preparative purifications were done with the PelB-GFPNb, which was chosen arbitrarily.

The purification yield for periplasmic and medium purification was considerably lower than cytoplasmic purifications, because the amount of functional protein is limited by the periplasmic export machinery rather than by the translation machinery. Thus, optimising yield is critical for these methods of purification. Various conditions were investigated to optimise yield as shown in **Figure 3.9**. It was found that in 50 mL cultures using LB medium did not produce sufficient Nb to be detectable by SDS-PAGE and Coomassie stain, whereas, M9 minimal medium cultures of the same size did (**Figure 3.9A**). Further optimisation showed that when 1 L minimal medium cultures were induced at OD600 ~0.7-0.8, and moved from 37ºC to a lower temperature for different lengths of time, approximately 5 hours of culture at 16ºC yielded the most protein after His-tag purifications and SDS-PAGE analysis, as shown in **Figure 3.9B.**

SDS-PAGE revealed two bands were observed at approximately 14 and 26-31 KDa. Reducing the GFPNb with 10 mM dithiothreitol (DTT) and re-analysis by SDS-PAGE, as shown in **Figure 3.9C,** almost entirely removed the higher MW band. Therefore, the higher band was probably a disulphide linked dimer. This conclusion was later confirmed by MALDI-MS analysis. The responsible residue was found to be the N-terminal Cys.

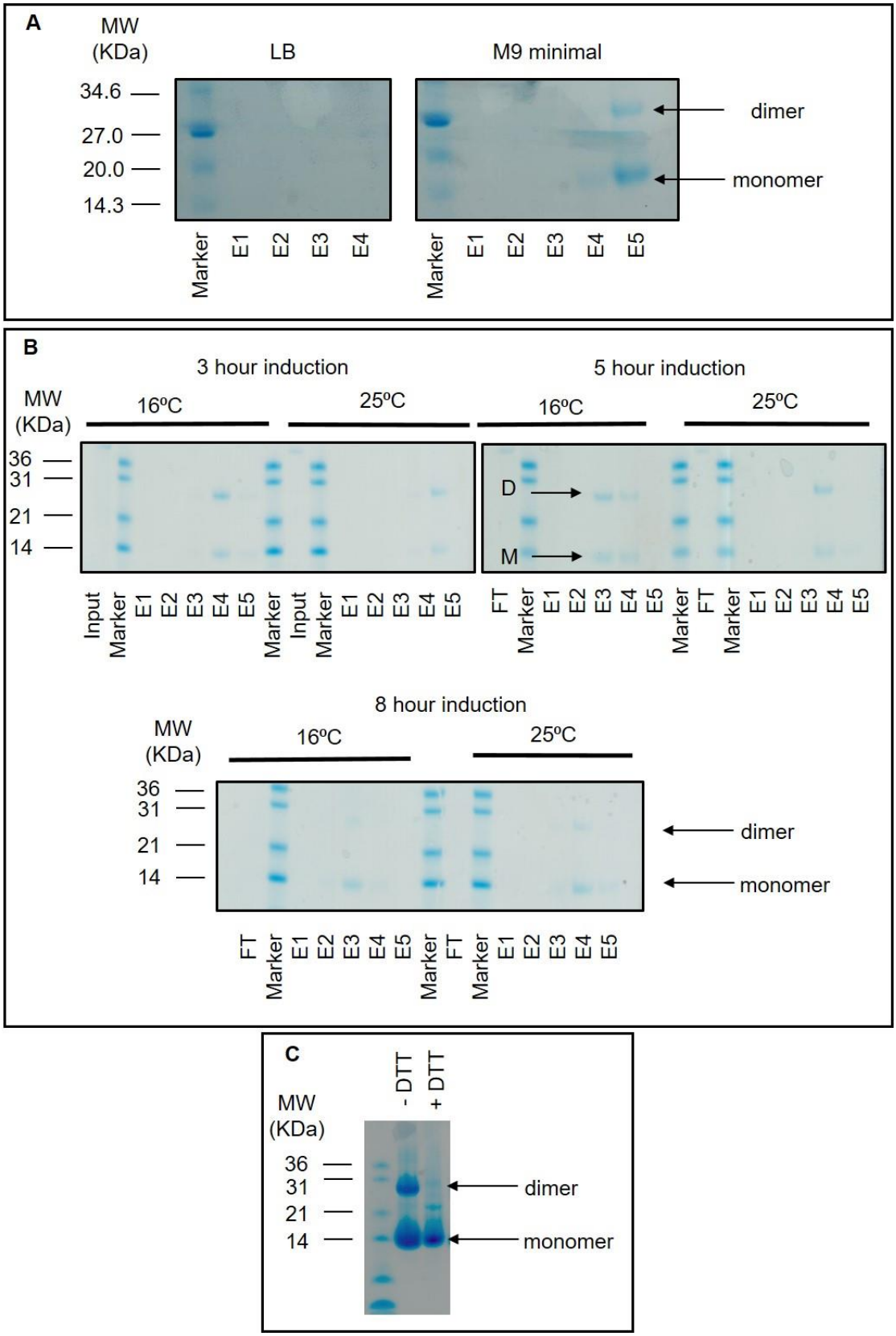


Figure 3.9 – Optimising GFPNb expression in E. coli strain LEMO21. A *SDS-PAGE comparison of expression using different culture media.* ***B*** *SDS-PAGE comparison of Nb yield after different induction temperatures and lengths of time.* ***C*** *SDS-PAGE gel with concentrated GFPNb after Ni-NTA purification, with and without 10 mM DTT reducing agent. (Dimer and Monomer, or D and M are indicated with arrows). 'E1' denotes 'elution fraction 1' and so on, FT indicates the flow-through from the nickel-nitrilotriacetic acid (NiNTA) column.*

3.2.2.3 Expression and Ni-NTA purification

When expressed in the LEMO21 *E. coli* strain the GFPNb was secreted to the periplasm and into the media using the optimised conditions described in **section 6.3**. It was purified via its C-terminal hexa-His-tag and analysed by non-reducing SDS-PAGE, as shown in **Figure 3.10A**. The gel showed protein eluted from a Ni-NTA gravity column in good yield, with two bands at approximately 14 KDa and 26-31 KDa, representing the monomer and disulphide linked dimer. Thus, the expression and Ni-NTA purification were successful.

3.2.2.4 MALDI-MS

It was important to confirm that the N-terminal Cys residue was present and available for NCL linkage to the dual labelled SBz-peptide. SDS-PAGE is not precise enough to determine this, so the Ni-NTA purified Nb was analysed by MALDI-MS.

The results of MALDI-MS showed experimental masses of 13,770; 13782.2 Da and 13796.7, with an expected mass of the monomeric Nb cleaved before C23 (PelB-GFPNb) of 13769.22 Da. This indicated that the correct product with the Cys free at the N-terminus was present. However, the other experimental masses did not correspond to either the correct mass nor to signal peptidase cleavage at other points. They were suspected to be post-translational modifications of the N-terminal Cys sulphur atom, because when the dimeric species was studied by MALDI-MS, no modifications were observed. This result meant that to ensure all the protein present is capable of NCL, it needed further purification.

3.2.2.5 Gel Filtration

One potential method was to use gel filtration chromatography to select for the dimeric species and then reduce that before NCL. As described in **section 6.3.5**, a 24 or 320 mL S75 column was used to separate the low MW species and the eluted fractions were analysed by non-reducing SDS-PAGE. The A280 curve is shown in **Figure 3.10B** and the gel in **Figure 3.10C**. The A280 trace shows that there were two distinct peaks, suggestive of successful separation, the dimer was expected to be the first peak, as it is larger. The corresponding elution fractions on the gel in **Figure 3.10C** (D10-D12) show a larger amount of dimer band than monomer. Conversely the later peak shows bands that contain more monomer. This indicates that separation of monomer and dimer was successful. However, there was likely to be some equilibration between the two states after gel filtration and before SDS-PAGE analysis.

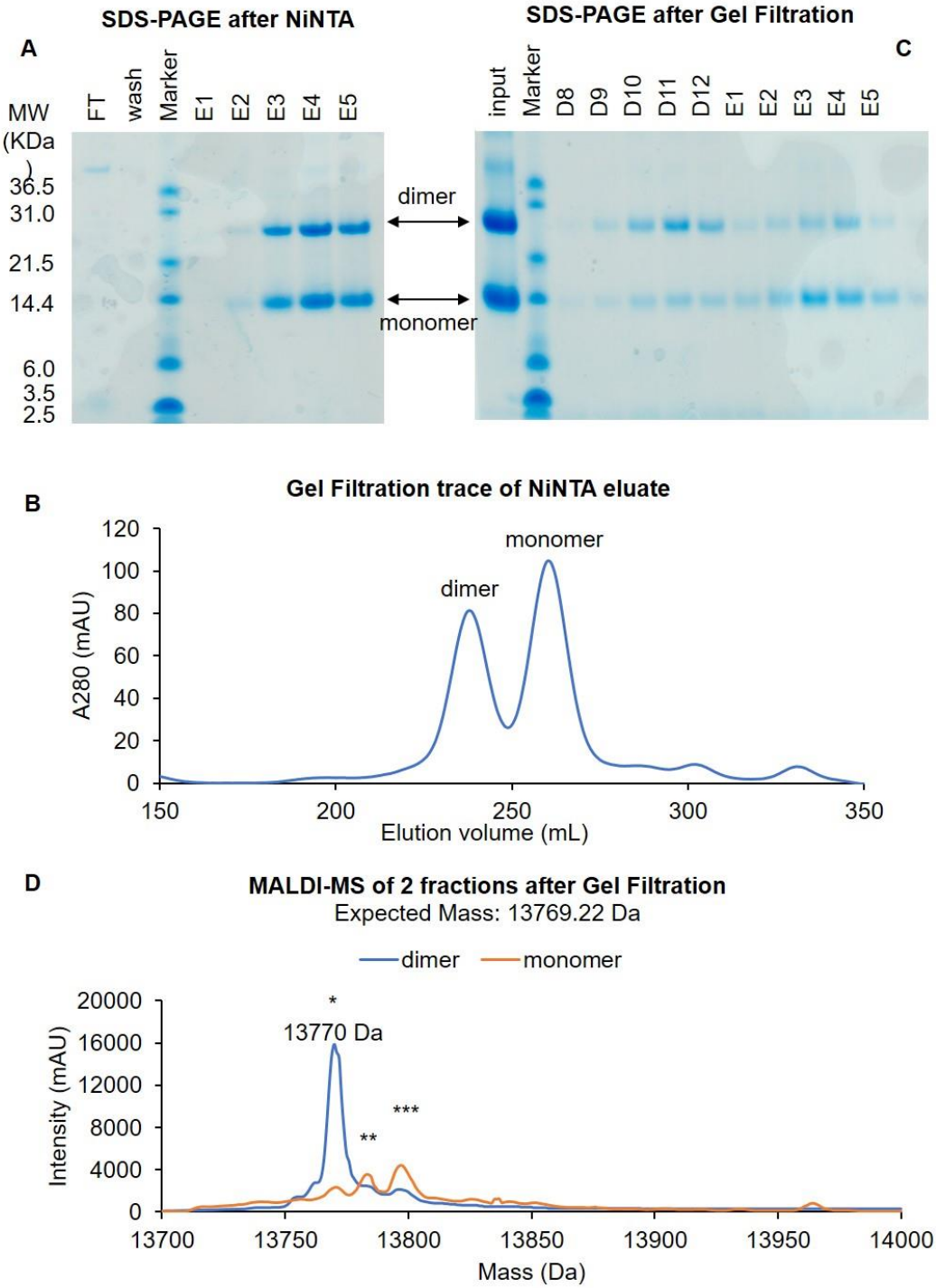


Figure 3.10 – Purification of dimeric GFPNb. A *Non-reducing SDS-PAGE gel of NiNTA purification shows monomer and dimer bands.* ***B*** *Gel Filtration A280 curve of Nb, on a S200 column facilitates separation of the monomer and dimer.* ***C*** *The non-reducing SDS-PAGE gel of the gel filtration fractions across the two dominant peaks.* ***D*** *The intact MALDI-MS results for the separated monomer (orange) and dimer (blue) peaks, with masses at (*) 13,770 Da, the expected mass of the monomeric Nb; (**) at 13782.2 Da and (***) at 13796.7, which represent chemically modified forms of the monomeric Nb.*

3.2.2.6 MALDI-MS analysis of monomer and dimer

To determine whether the N-terminal Cys of the dimer was free, unmodified and capable of NCL, MALDI-MS was carried out on the separated monomer and dimer fractions. The results are shown in **Figure 3.10D**. The monomer sample showed

multiple peaks, but the dimer only sample mainly contained a single species, with mass equal to 13770.0 Da, within error of the expected mass of 13769.22 Da. This showed that the dimer population had a protected N-terminal Cys, which was likely functional for NCL. The monomer population on the other hand had several unexplained peaks in addition to the expected one, which were likely the result of post-translational modification, which made it useless for NCL.

3.2.2.7 NCL

The final step of STORM probe synthesis was to link the labelled SBz-peptide from **section 3.1** to the purified GFPNb and to confirm its identity. The desired product is referred to as GFPNbCy3Cy5 and its structure and expected mass are shown in **Figure 3.11C**. The NCL reaction was carried out according to a previously published protocol (Bartke et al. 2010), as described in **section 6.6.4**. NCL was followed by Gel Filtration on a NAP-5 column to separate the unreacted SBz-peptide from the labelled protein and the elution fractions were analysed by non-reducing SDS-PAGE and imaged for fluorescence as shown in **Figure 3.11A** and **B**. This showed that the concentration of protein was too low to be observed by Coomassie stain, but two fluorescent Cy5 bands were detected at a MW of approximately 14 KDa. This indicated that the protein has been covalently linked to the dual labelled SBz-peptide. To confirm labelling, the final product was also analysed by MALDI-MS (**Figure 3.11D**). A mass was observed at 15503.6 Da within error margins of the expected mass 15502.37 Da for GFPNbCy3Cy5 shown in **Figure 3.11C** and a yield of approximately 90%. This meant that ~10% of the GFPNb remained unlabelled, which needed to be taken into consideration for later quantification experiments. While better resolving methods such as ion exchange chromatography were investigated to separate peptide labelled Nb from unlabelled, none of these approaches were successful.

Overall, a GFPNb had been successfully expressed and purified for NCL, and linked to the SBz-peptide-Cy3-sulfo-Cy5 to form GFPNbCy3Cy5. The NCL reaction with GFPNb was also carried out using SBz-peptide-Alexa488-sulfo-Cy5 with similarly successful results (not shown).

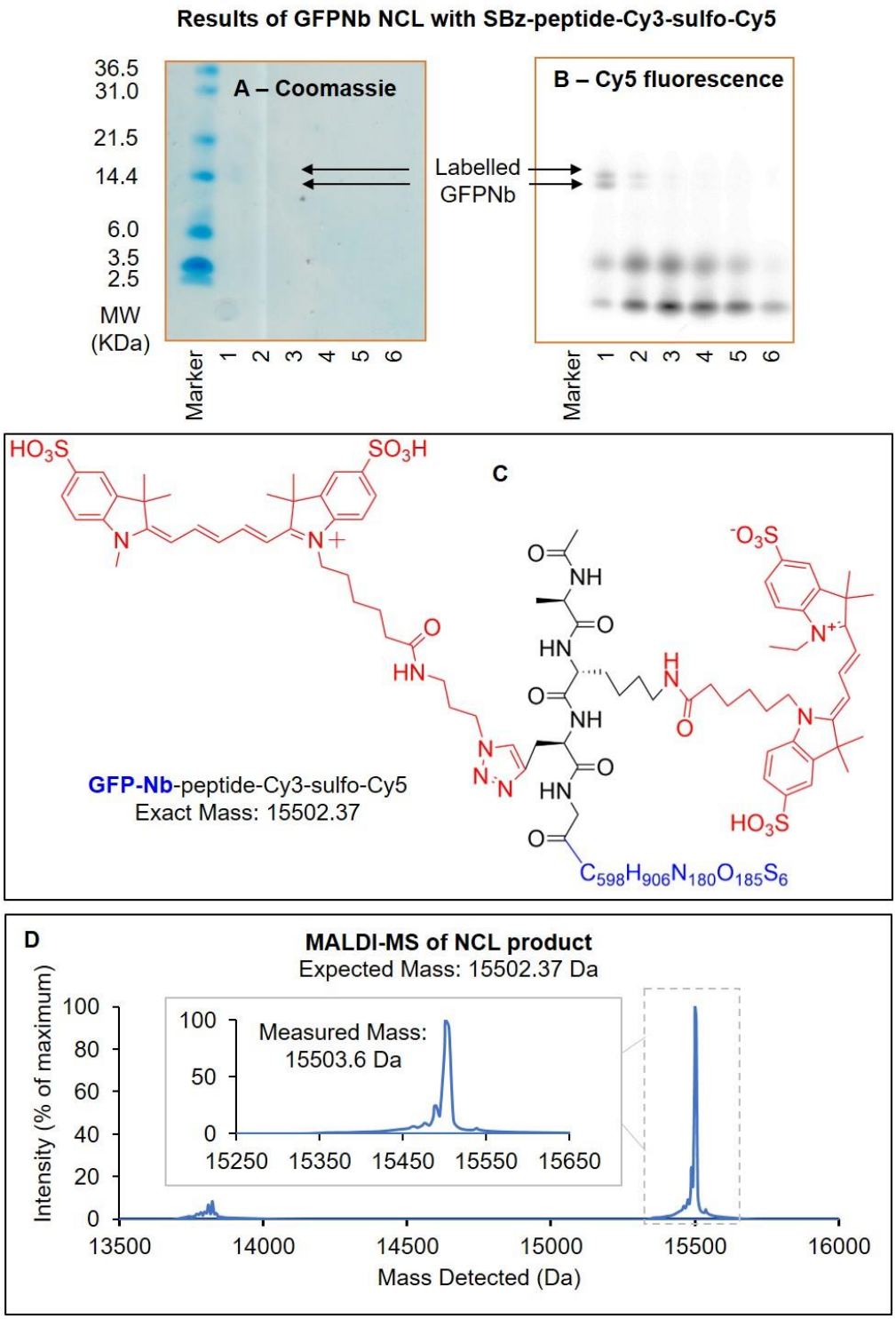


Figure 3.11 – NCL successfully links together GFPNb and doubly labelled SBz-peptide. ***A*** *and* ***B*** *The same non-reducing SDS-PAGE gel imaged using* ***A*** *Coomassie stain and* ***B*** *Cy5 fluorescence.* ***C*** *The structure of the desired GFPNbCy3Cy5 with the dyes highlighted in red and he GFPNb simplified and highlighted in blue.* ***D*** *MALDI-MS shows that the product from NCL has a mass of 15503.6 Da consistent with the expected masses of 15504.3 or 15502.3 when ionization is dominated by the 4' ammonium or protonation, respectively.*

3.2.2.8 Testing the labelled GFPNb

To test whether this dual labelled GFPNb construct was a viable STORM probe, it was necessary to test its ability to bind their respective antigens, because it may not have

folded correctly and the addition of the dual labelled peptide may have hampered the epitope binding site. GFPNbCy3Cy5 binding was tested using the *S. pombe* centromeric protein Cnp1 (CENP-A), which had already been labelled with GFP (Takayama et al. 2008).

Cultures of this strain were grown, and made into spheroblasts, permeabilised and labelled as described in **section 6.8.1**. They were imaged for Cy5 signal at low 641 laser intensity to prevent Cy5 blinking, and the images compared to ensemble images of GFP. Good labelling was expected to produce co-localised Cy5 and GFP in one of two foci in the middle of the cell.

For GFP-Cnp1 plus GFPNbCy3Cy5 the results for one representative field of view are shown in **Figure 3.12**. All the cells in the white light image have GFP foci when illuminated with 488 nm laser light, as expected, although some are dimmer, due to imperfect laser illumination. The same field of view imaged with 561 and 641 lasers to detect Cy3 and Cy5 fluorescence, which show bright foci in most, although not all of the cells. **Figure 3.12E-F** show that for one of these cells the fluorescence signal for all three channels increases at the same position on a cross-section through one cell – a strong indication of co-localisation of the foci in different channels. Taken together these data show that the Cy3 and Cy5 fluorescence co-localise with the GFP-Cnp1 fluorescence at the centromere, strong evidence of good labelling, meaning the GFPCy3Cy5 construct is functional for target binding, and a good candidate for a new STORM probe.

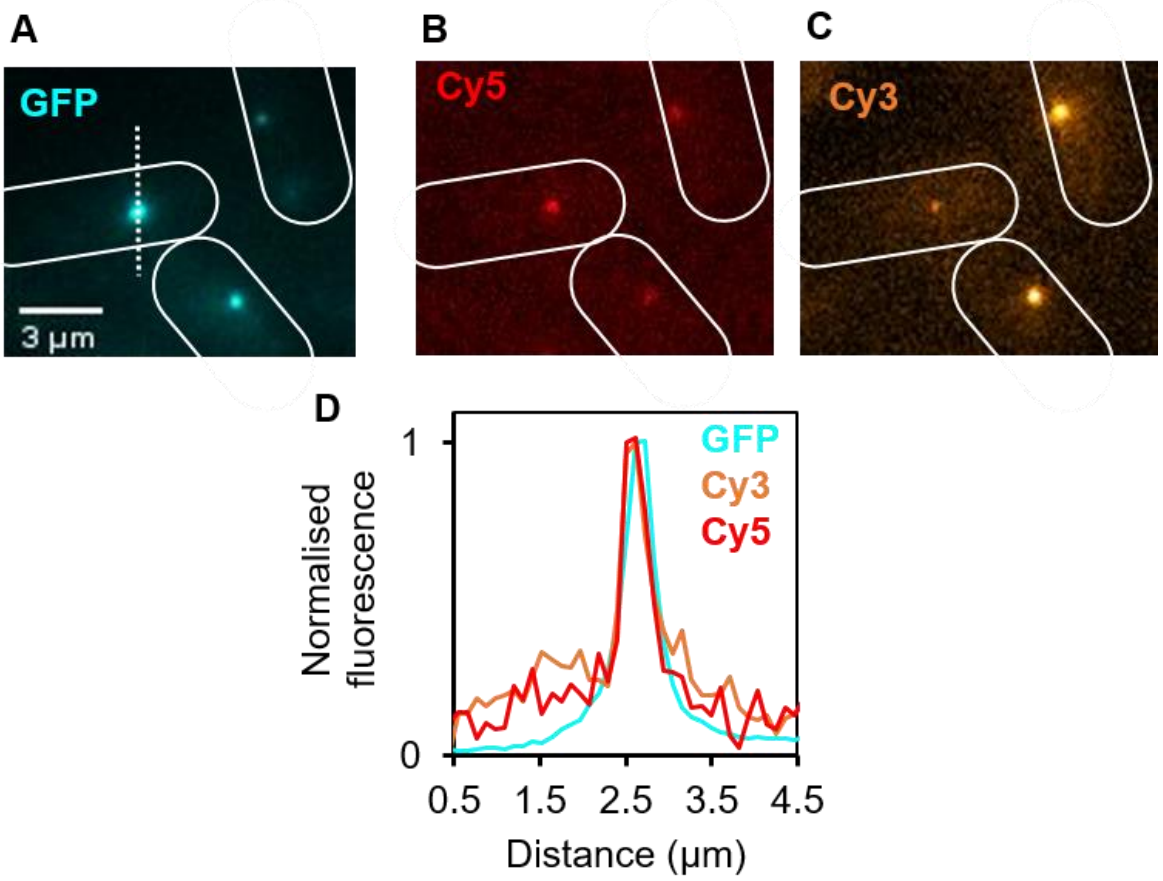


Figure 3.12 – GFPNbCy3Cy5 successfully labels GFP-Cnp1 in S. pombe cells. ***A-D*** *show the same field of view containing 3 cells;* ***A*** *GFP,* ***B*** *Cy5****, and C*** *Cy3.* ***D*** *Fluorescence profile for all three colours along the dotted line indicated in* ***A****. Fluorescence was normalised to maximum.*

3.2.3 An anti-mCherry-Nb for periplasmic expression and NCL

Successful expression, purification and labelling target binding of the GFPNb construct potentially allowed for single colour STORM imaging. However, two different activator-emitter constructs with different targeting modules are required to attempt dual colour STORM imaging. Therefore, we also attempted similar purification, expression and labelling with an anti-mCherry-Nb (hereafter mCherryNb).

An mCherryNb, LaM-4 from Fridy et al (2014), was selected from the six candidates they characterised, based upon its high affinity for mCherry and low cross-reactivity for GFP (Fridy et al. 2014). It was also adapted for periplasmic expression in *E. coli*, His-tag purification and subsequent NCL. It was also based upon the successfully expressed and NCL linked GFPNb designed and expressed in **section 3.2.2**. The construct had an N-terminal DsbA signal sequence with a subsequent CysGlu to optimise signal peptidase cleavage, as shown in **Figure 3.13** aligned with the original LaM-4 sequence (Fridy et al. 2014) and DsbA-GFPNb sequence for comparison.

```
DsbA-GFPNb          MKKIWLALAGLVLAFSASACEVQLVESGGALVQPGGSLRL 40
DsbA-mCherryNb      MKKIWLALAGLVLAFSASACEVQLVESGGSLVQPGGSLRL 40
mCherryNb           ------------------MAQVQLVESGGSLVQPGGSLRL 22
                                       .:********:**********

DsbA-GFPNb          SCAASGFPVNRYSMRWYRQAPGKEREWVAGMSSAGDRSSY 80
DsbA-mCherryNb      SCAASGRFAESSSMGWFRQAPGKEREFVAAISWSGGATNY 80
mCherryNb           SCAASGRFAESSSMGWFRQAPGKEREFVAAISWSGGATNY 62
                    ******  .:  ** *:**********:**.:* :*. :.*

DsbA-GFPNb          EDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYYCNVNV 120
DsbA-mCherryNb      ADSAKGRFTLSRDNTKNTVYLQMNSLKPDDTAVYYCAANL 120
mCherryNb           ADSAKGRFTLSRDNTKNTVYLQMNSLKPDDTAVYYCAANL 102
                     **.*****:***:::************:******* .*:

DsbA-GFPNb          G---------FEYWGQGTQVTVSSKHHHHHH 142
DsbA-mCherryNb      GNYISSNQRLYGYWGQGTQVTVSSKHHHHHH 151
mCherryNb           GNYISSNQRLYGYWGQGTQVTVSSPFT---- 129
                    *         : ************ .
```

Figure 3.13 – Sequence alignments of the nanobodies: *GFPNb and mCherryNb both with the DsbA signal sequence and hexa-His-tag modifications compared to the original mCherryNb* (Fridy et al. 2014*; LaM-4). The comparison underneath denotes similarity between the sequences; * an identical residue, : a different residue with strongly similar chemical properties, whereas . denotes a different residue with weaker chemical similarity. Red regions indicate the 3 CDRs, and the N-terminal signal sequences are shown in orange as well as the subsequence CysGlu, in blue, which are revealed after signal peptidase cleavage.*

Apart from the essential changes required for periplasmic export, purification and NCL, several small modifications to the N- and C-termini of LaM-4 were adapted to mirror the GFPNb as closely as possible because it had expressed, transported and purified

very effectively, thus the 3 N-terminal residues of LaM-4 were removed and replaced by the crucial CysGlu to maximise likelihood of effective cleavage by the periplasmic signal peptidase, as suggested by Choo et al (K. H. Choo, Tong, and Ranganathan 2007). Similarly, the 3 C-terminal residues of LaM-4 were removed and replaced with the hexa-His-tag successfully used to purification with the GFPNb.

Notably, this construct is slightly larger than the GFPNb, with a predicted mass of 14406.90 Da after signal peptidase cleavage, compared to 13769.22 Da for the GFPNb.

3.2.3.1 Expression and purification

This mCherryNb construct was expressed in a similar way to the GFPNb using the LEMO21 *E. coli* strain and IPTG induction at 16ºC in M9 minimal medium to secrete the protein into the periplasm for folding and signal peptidase cleavage of the N-terminal DsbA signal sequence.

The expression and purification of the protein was analysed by SDS-PAGE and MALDI-MS **(Figure 3.14A)**. Both methods revealed that both monomer and dimer species were slightly truncated at the C-terminus, since the N-terminal Cys is required for dimerisation. Separation of the monomer and dimer was performed by gel filtration as shown in **Figure 3.14B**, and analysed by SDS-PAGE (**Figure 3.14C**). The results show that the dimer was successfully separated by gel filtration. MALDI-MS was then used to asses if the N-terminal Cys is capable of NCL (**Figure 3.14D**). Unfortunately, it showed that whilst the expected mass of 14406.90 Da peak was present, most of the detected ions were not at the expected mass. The 12829 Da peak was determined to be the result of a contamination during the MS process itself as it was observed during multiple runs of unrelated samples, so does not reflect any property of the mCherryNb. The additional higher than expected masses of +30-100 Da in **Figure 3.14D**, proved more difficult to remove. It was thought that masses with more than expected mass could result from oxidation of the N-terminal Cys, so reduction was attempted with 5 mM DTT at 37 °C for 30 min. This reduced the size of the 14436.4 and 14470.7 Da peaks but also reduced the size of the correct 14406.1 Da peak, so it was judged ineffectual. A solution for this sub-optimal result was not found owing to time pressure. It was hoped that the modifications were not to the N-terminal Cys itself but elsewhere on the protein, but they eventually prevented use of the mCherryNb for imaging.

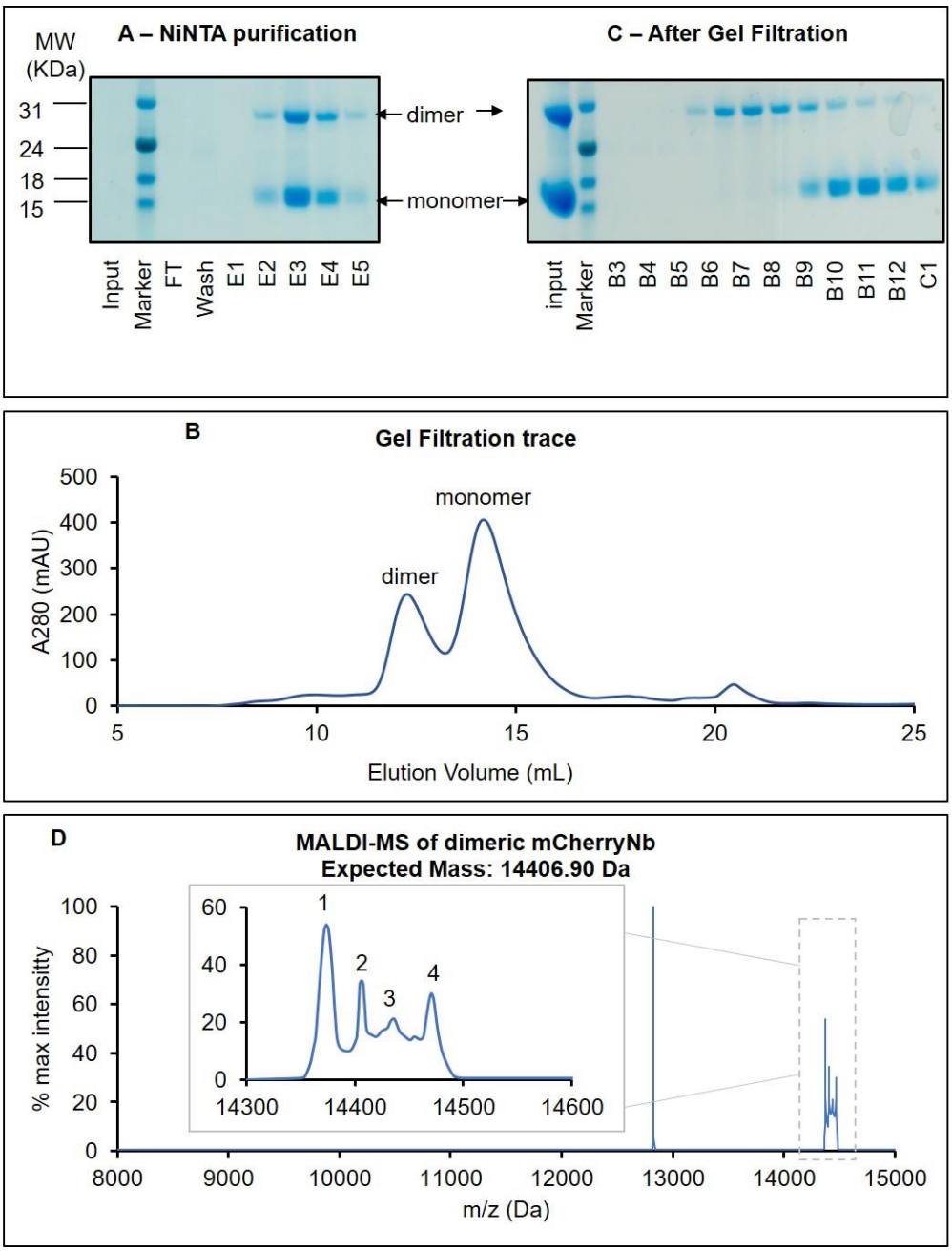


Figure 3.14 – Purification of dimeric mCherryNb. A *Non-reducing SDS-PAGE gel of NiNTA purification shows monomer and dimer bands.* ***B*** *Gel Filtration A280 curve of Nb on a S75 column facilitates separation of the monomer and dimer.* ***C*** *Non-reducing SDS-PAGE gel of the gel filtration fractions across the two dominant peaks.* ***D*** *The intact MALDI-MS results for the separated dimer factions (BX-Y from* ***C****), with major mass peaks at 1 – 14373.736 Da, 2 – 14406.083 Da, 3 – 14436.428 and 4 – 14470.679 Da plus a large peak at 12829.906.*

3.2.3.2 NCL

The final step in modifying this mCherryNb for STORM probe was to link the dual labelled SBz-peptide via NCL, and confirm that the labelling worked. NCL was carried out as described in **section 6.6.4**. The reaction product was purified using a NAP-5 column and analysed by SDS-PAGE, UV-visible absorption spectroscopy and MALDI-MS. The SDS-PAGE results are displayed in **Figure 3.15A-C**; they showed a pair of

bands at approximately 18 KDa in the Coomassie stained gel as well as when imaged for Cy3 and Cy5 fluorescence, which indicated that the protein was successfully labelled. There was also a fluorescent band below the smallest 8 KDa marker, which did not stain with Coomassie. This was probably the unreacted SBz-peptide-Cy3-sulfo-Cy5. The absorption spectrum of the product, shown in **Figure 3.15D** also supported labelling as the product absorbed at 280 nm, as well as at ~550 nm and ~650 nm, as expected for a protein containing Trp, Cy3 and Cy5, respectively.

MALDI-MS was used to confirm labelling and examine the extent of the reaction, as shown in **Figure 3.15E**. The mass spectrum showed the presence of at least three distinct species, each with several peaks that are much closer in mass. The set of peaks at ~14464.78 Da contains major masses of 14405-14569 Da, which corresponds to the unlabelled Nb (expected mass 14406.90) and the covalent modifications also observed in **Figure 3.14D**. The peaks at 16135.477 & 17870.960 Da comprise just one major mass peak, which correspond to the mCherryNb with one and two dual labelled peptides attached via the N-terminal Cys (singly labelled, referred to as mCherryNbCy3Cy5, and doubly labelled, mCherry(NbCy3Cy5)$_2$), respectively, which have expected masses of 16129.58 and 17877 Da. There is also a peak at 8936.897 Da, which is the doubly charged ion of the 17870.96 Da peak. Whilst MS is not quantitative because different ions fly differently, the general indication seems to be that more than 1/3 of the mCherryNb is labelled with two peptides (therefore four dyes), whereas, less than 1/3 is labelled with a single STORM peptide, and less than 1/3 remains unlabelled. The double-labelling was problematic, because it prevented reliable quantitative measurements. It was probably caused by the N-terminal Cys linking to one SBz-peptide via its primary amine and another via its sulphydryl group, thus it should be easily removed by addition of a reducing agent such as DTT, but this experiment was not carried out due to time pressure.

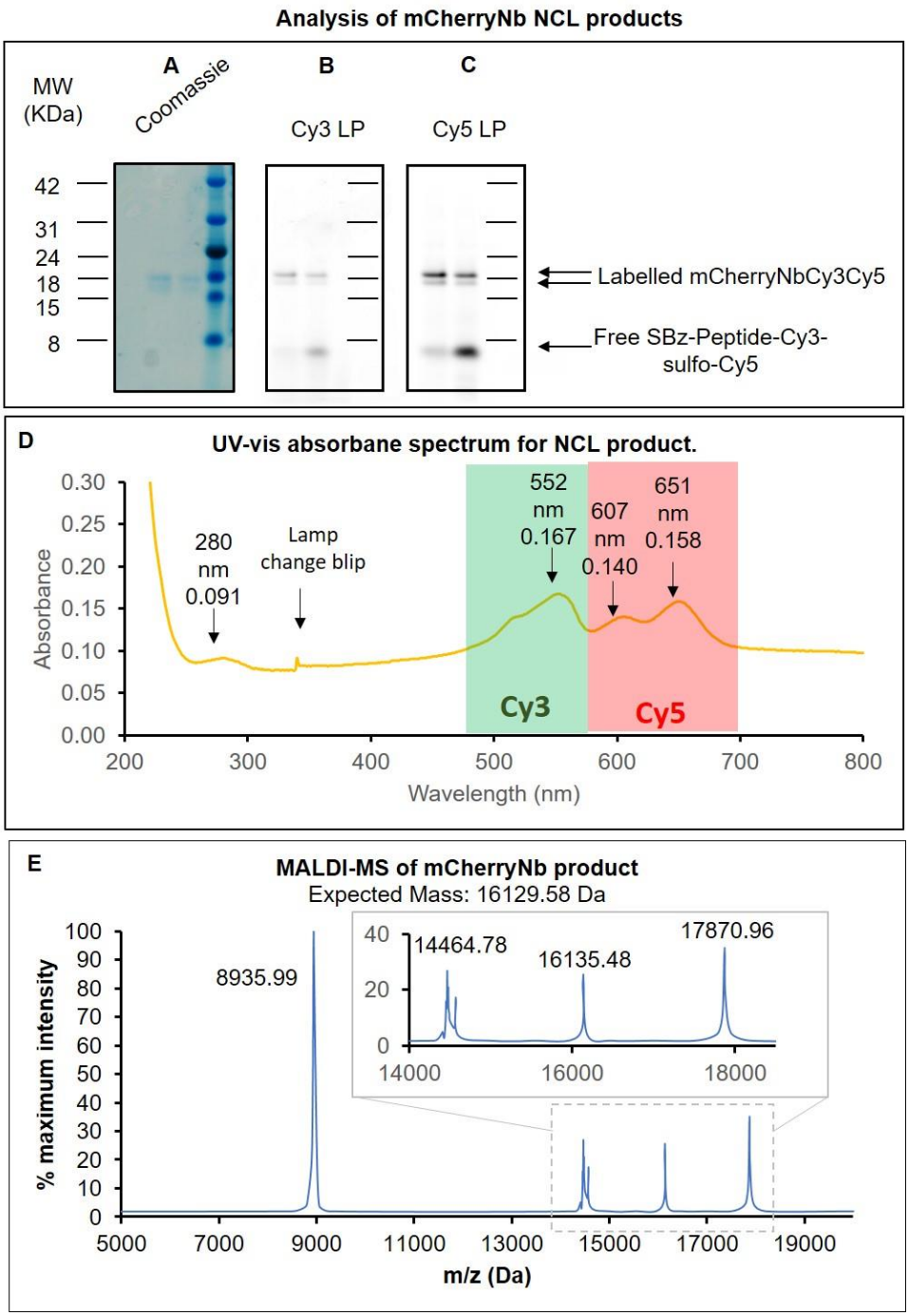


Figure 3.15 – NCL successfully links mCherryNb and dual labelled peptide. A-C *The same non-reducing SDS-PAGE gel of Nb imaged using* ***A*** *Coomassie stain* ***B*** *Cy3 fluorescence and C using Cy5 fluorescence. 2 bands for mCherryNbCy3Cy5 and 1 band for the unreacted Peptide-Cy3-sulfo-Cy5.* ***D*** *UV-visible absorbance spectrum of the protein bands from* ***A****.* ***E*** *MALDI-MS trace with peaks for the dominant masses. The expected product has a mass of 16129.58 Da. With an inset zoomed in on the three important peaks.*

These results show at least partially successful labelling of the mCherryNb with SBz-peptide-Cy3-sulfo-Cy5 via NCL and confirmed by MALDI-MS. However, the high degree of unlabelled, doubly labelled and unidentified covalent modifications meant that this reagent was not good enough to use for further imaging work, so this construct was abandoned with insufficient time to optimise the purification and labelling.

3.3 GFPNbCy3Cy5 for preliminary STORM imaging of GFP-Cnp1

Section 3.2.2 presented good evidence that we could synthesise an epitope binding GFPNb labelled stoichiometrically with Cy3 and Cy5 via an SBz-peptide. We therefore proceeded to image the construct under conditions suitable for Cy5 blinking to show the construct could be useful for SMLM imaging in cells. The protocols and analyses are described in detail in **section 6.8.4** and a summary is shown in **Figure 3.16A**.

Centromeric clusters of Cy5 localisations were used to estimate the number of GFP-Cnp1 at the centromere, since previous work had investigated the number of Eos2-Cnp1 in *S. pombe* cells (Lando et al. 2012), it was thought possible to roughly measure the number of GFPNbCy3Cy5 and compare to the published results as a preliminary test. Previous counting methods have used numbers of blinks (Hummer, Fricke, and Heilemann 2016; Karathanasis et al. 2017; Krüger et al. 2017; Lando et al. 2012) and numbers of localisations (Ricci et al. 2015; Zanacchi et al. 2017) to count fluorophores, however these were not thought to be effective approaches of counting GFPNbCy3Cy5 in the context of Cnp1, because both would be affected by the high density of GFPNbCy3Cy5 at the centromere cluster and large number of blinks Cy5 before bleaching, and so analysis was likely to underestimate both number of blinks and localisations, especially when multiple Cy5 could have been activated simultaneously. Instead, we opted to measure total intensity detected from the fluorophores, which was thought not to be affected by high Cnp1 density, and divide that by the average total number of photons produced by a GFPNbCy3Cy5 molecule on a SM surface (378.9 photons per GFPNbCy3Cy5, this was from data discussed later in **section 4.2.7**). The results are shown in **Figure 3.16B-E**.

The data showed that Cy5 could be made to blink in and allow SR image reconstruction (**Figure 3.16C**), therefore, the GFPNbCy3Cy5 was a shown to be a useful SMLM tool. Cy5 was also readily activated by manual 561 nm activation, which was good evidence that the STORM probe was also functional, although we could not, at this stage, discount the possibility that Cy5 absorbed the 561 nm light directly causing it to re-enter the on-state.

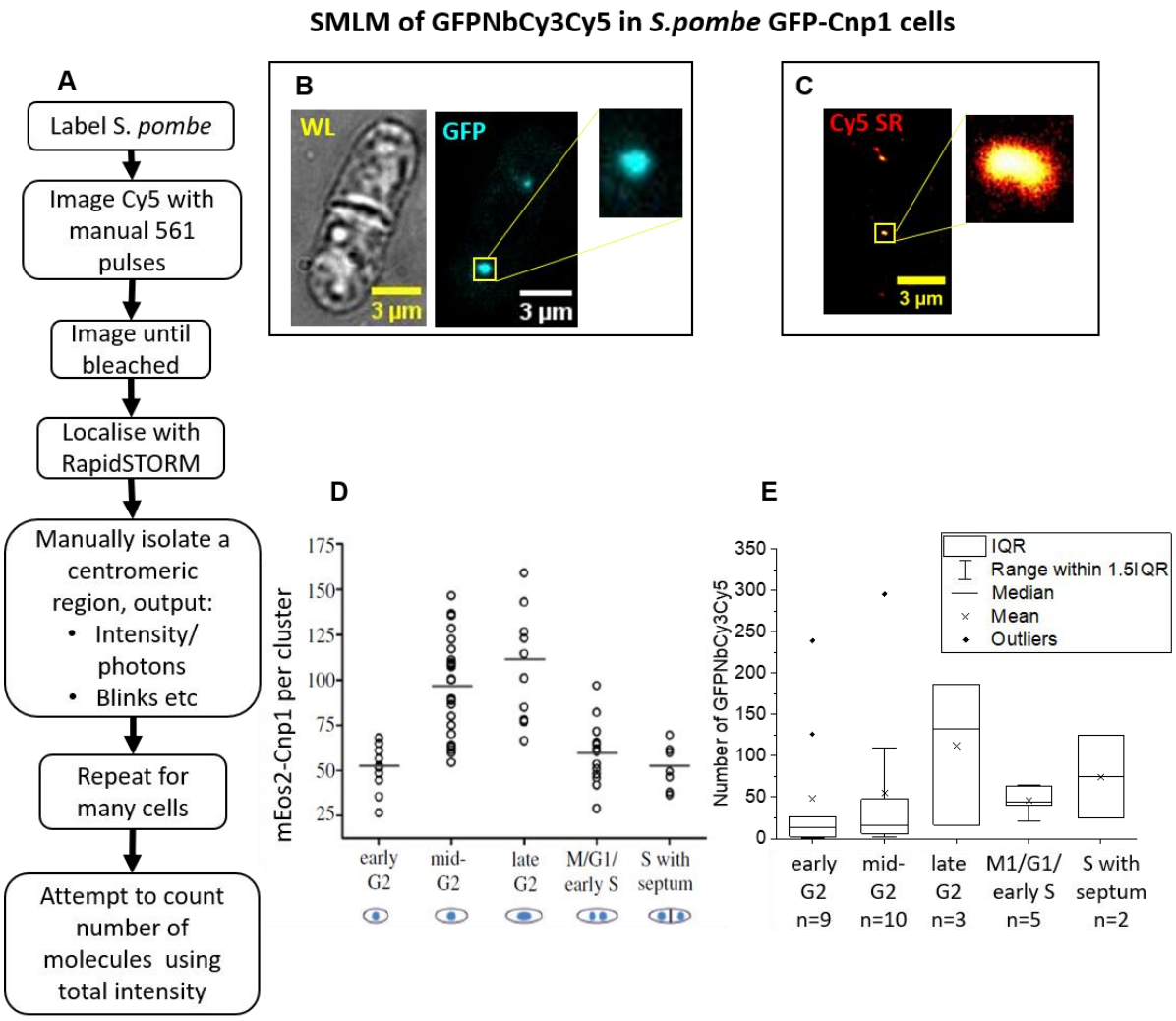


Figure 3.16 - Premilitary STORM data acquisition in S. pombe cells. A *Methodology scheme for the experiment.* ***B*** *Wight light (WL) and GFP images of an S. pombe cell highlight the centromere, with a zoom in on the centromere inset* ***C*** *Super-resolution image of the same S. pombe cell after imaging, localisation and image reconstruction. A zoom in of the centromere Is inset.* ***D*** *Previously published box plot showing number of mEos2-Cnp1 molecules per centromere* (adapted from Lando et al. 2012, Figure 4A). ***E*** *Box and whisker plot of number of GFPNbCy3Cy5 detected in this experiment calculated by number of photons, the data for a total of 29 cells was used.*

Whilst several technical problems were encountered, data for ~ 50 cells were initially collected over several weeks, and of these 29 videos were judged sufficient quality to analyse for numbers of molecules, as shown in **Figure 3.16E**. Other videos were discounted due to poor laser alignment, Z-drift, insufficient blinking and apparently short total on-times. The box plot showed some qualitative similarity to the numbers of molecules previously observed by Lando et al (2012; **Figure 3.16D**), because the general increase in mean numbers of GFPNbCy3Cy5 from early- to late G2 phase and subsequent decrease was evident. However, the absolute number of molecules and statistics were not as similar as we had hoped, for example the mean number of GFPNbCy3Cy5 detected in early G2 was 48 with a range of 238 in our dataset

compared to the previous data mean ~ 50 and range ~50. Clearly the GFPNbCy3Cy5 produced more variation that the mEos2 data, but it was difficult to assign the cause of this to any one factor due to several technical limitations as well as the inherent day-to-day variation of the setup. Nevertheless, there was a degree of similarity in the number of GFP-Cnp1 to mEos2-Cnp1, which indicated several key points: firstly, GFPNbCy3Cy5 appeared to label most of the GFP-Cnp1 in the cells, secondly, SMLM imaging was possible in the cells, with 561 nm STORM activation a useful option in this context. This was important because *S. pombe* have seemed difficult to use with Halo- and SNAP-dyes for *d*STORM due to lack of blinking so PA-FPs have been preferred (Endesfelder et al, unpublished).

Imaging GFPNbCy3Cy5 in cells was a significant step to show it could be a useful SMLM tool and may be comparable to other quantitative methods like mEos2. Further work counting GFP-Cnp1 was not pursued due to microscope availability and large amounts of time taken to collect sufficient data. Future approaches could use more stable setups, fiducial markers to correct for XY drift, and possibly an asymmetric lens to allow multiple centromeres to imaged simultaneously and reduce total imaging time. Additional methods could use cell cycle synchronisation to ensure sufficient numbers of cells from each cell cycle phase are collected.

3.4 Summary

The aims of the work in this chapter were to synthesise a dual labelled STORM probe and link one copy to a functional protein targeting domain, based upon either an scFv or a Nb. This work involved optimising the chemospecific NHS and click chemistry reactions to produce a dual labelled peptide, as well as the design, optimisation, expression and purification of potential scFvs and Nbs to link that peptide to, as well as testing the produced constructs for epitope binding.

The SBz-peptide was successfully dual labelled with Cy3 and sulfo-Cy5 as well as Alexa488 and sulfo-Cy5 dyes in close proximity, which are thought to be suitable for STORM. The labelled peptides were purified by RP-HPLC and chemical structure confirmed by MALDI-MS.

Expression of periplasmic scFvs against the FLAG and c-Myc epitopes proved unsuccessful. Whilst expression was detected by Western Blot, they were only ever present in the insoluble fraction of the cellular pellet, this was indicative of aggregation.

In contrast, two Nbs against GFP and mCherry were successfully expressed in the periplasm of LEMO21 *E. coli* cells. Initial purifications were from LB medium, but optimisation of these conditions showed that the use of M9 minimal medium and induction at 16ºC for 3-5 hours produced a higher yield. SDS-PAGE analysis revealed that both Nbs can form N-terminal disulphide linked dimers. MALDI-MS showed that both GFPNb and mCherryNb were post-translationally modified, with multiple mass peaks above that of the expected mass. In the GFPNb case, MALDI-MS after separation of the monomer and dimer fractions by gel filtration showed that the dimer contained little modification. Gel filtration was therefore carried out before NCL, to ensure as much functional N-terminal Cys as possible was available. For the mCherryNb it was not possible to obtain an unmodified sample, for reasons that are currently not known. Covalent modification occurs at some point during expression or purification, which is not removed by the DTT reduction that was attempted. This indicated that the modifications are probably not caused by oxidation of the N-terminal Cys sulphhydryl group, but may be by modifications of the N-terminal amine group, which would also preclude the acyl shift step in NCL.

NCL was successfully used to link the chemically synthesised SBz-peptide-Cy3-sulfo-Cy5 to both Nbs, and seemed to work robustly. Correct labelling of the GFPNb was confirmed by MALDI-MS. Although the amount of labelling was variable between batches, a batch was obtained with over 90% singly labelled GFPNb (**Figure 3.11**). Such little material is required for SMLM, this one well labelled batch was sufficient for subsequent experiments.

GFPNbCy3Cy5 labelled GFP-Cnp1 in *S. pombe* with high specificity since microscope images showed low Cy3 and Cy5 fluorescence in most of the cell apart from a very high intensity focus at the centromere, which co-localised with GFP fluorescence. This construct was then used to show that SMLM imaging using 561 nm activation was possible and produced SR images and reasonable although variable numbers of GFP-Cnp1 at the centromere cluster.

3.5 Discussion

The successful expression of two Nbs in LEMO21 *E. coli* and the failure to express two scFvs was not particularly surprising because the Nb sequences were taken directly from published work which required good expression for structural and functional studies (Fridy et al, 2014; Kubala et al, 2010). The scFvs on the other hand, were

designed based on Fab fragments (Rooslid et al, 2006; Hipbert et al, 2001; Shiweck et al, 1997) and an scFv scaffold (Kaufmann et al, 2002). Clearly rational design in this way cannot take account of all the factors needed for successful transport and folding of an scFv. There may be several reasons, however, it is well known that scFv development is done commercially by randomly combining V_H and V_L domains from polyclonal IgG antisera. Most combinations cannot bind their epitope and so are not selected by phage display and similar methods, so the likelihood of any one design expressing and folding correctly here was also low. Despite using constructs with optimised stability, the chances of successful folding were low and the constructs proved unable to transport to the periplasm. Aggregation in the cytoplasm, even at low expression rates, indicated that the constructs' hydrophobic regions led to aggregation before transport could occur. This may have been due to the *E. coli* expression system used here, which is naturally very fast compared to eukaryotic or mammalian expression systems which may have been more successful for these constructs, also with more chaperones aggregation may have been prevented.

Furthermore, different Abs develop different ligand affinities through somatic hypermutation and VDJ recombination during B-cell maturation. These processes produce insertion and deletion mutations in the F_V domain relative to other Abs as well as substitutions. The design process here only took account of small changes in structure between the anti-His-tag scFv and the anti-c-Myc/anti-FLAG Fab structures, mostly by substituting in residues in the CDRs. As shown in **Figure 3.13**, however, the CDRs vary in length between F_Vs and the anti-c-Myc V_H is 13 residues longer than either of the other two. Arranging the appropriate CDR residues onto a shorter scFv scaffold maybe the reason the anti-myc-scFv was unable to fold correctly, causing aggregation. The anti-FLAG F_V was more like the anti-His-tag F_V, at the same length, so may have been more likely to fold correctly, but the CDRs are in subtly different positions. These loops contain the residues for epitope binding, and so being out of position a little compared to the scaffold, may have caused the structural problems preventing expression. Furthermore, the residues in the loops are often hydrophobic to promote epitope binding via favourable exclusion of water, so a subtle imbalance of hydrophobic nature may have triggered unfolding of the anti-FLAG construct, but not be a problem in the anti-His-tag case.

The likelihood of expressing rationally designed scFvs is not straightforward, and this underlines the reasons that commercial production of scFvs and mAbs use phage

display and panning methods based on large libraries rather than designing them based on previous structures.

The Nbs on the other hand had no problem with aggregation, because they had previously been selected for bacterial expression. The work here to optimise them for periplasmic expression and NCL was untested but much more straight forward given previous knowledge about periplasmic secretion and signal peptidase selectivity (K. H. Choo, Tong, and Ranganathan 2007). The limiting aspects for this work were protein yield and N-terminal Cys post-translational modifications. When these Nbs were previously expressed for structural and functional purposes, they were expressed in the cytoplasm, which typically produces large yields, that approach was not used here because the N-terminal Cys for NCL was designed to be revealed by signal peptidase cleavage. Other proteases could have been used to reveal an N-terminal Cys, such as *Tobacco Etch Virus* nuclear-inclusion-a endopeptidase (TEV), but a suitable cleavage site had not been introduced in these constructs.

Periplasmic expression, was limited by the secretion machinery, presumably in part the SecYEG complex as well as the chaperones involved in delivery from the ribosome. This was probably why low temperature growth conditions as well as minimal culture medium helped increase the yield, because under these conditions cellular processes were likely to be much slower, including transcription and translation. This slowed the rate of protein production which likely suited the slow process of periplasmic export. Preliminary results in which the rhamnose tuneable LEMO21 expression strain were used, also hinted at this; adding rhamnose to indirectly inhibit RNA polymerase also increased Nb yield. It may have been possible to increase yield further by combining the periplasmic purification with the affinity chromatography directly from the medium, although the medium alone was sufficient as a source for this work.

NCL to link the dual labelled peptide to the Nbs with N-terminal Cys residues was also successful, and required little optimisation. Purification of unmodified Nb was eventually optimised in the GFPNb case, which was labelled in excess of 90%. The problem of post-translational modifications was frustrating, because it was not discussed in the literature and so potential solutions were not readily available. We discussed this problem with Medimmune and John McCafferty's group, who both express scFvs commercially. Neither group had any previous knowledge of the modifications observed here by MS, partly because they do not usually express

proteins with N-terminal Cys residues. This may be a unique problem, and the solution we came to in using the dimer may not be the optimal one.

Future work is needed to look at reducing post-translational modifications of the mCherryNb or separating the labelled and unlabelled constructs. This was briefly investigated by both ion exchange chromatography and RP-HPLC, but the Nb had very close to netural pI values and so proved difficult to separate even at very low and very high pHs. This could be especially useful since the GFPNb proved effective for labelling and STORM imaging, and a second construct with a second STORM pair would be useful to attempt multicolour imaging.

Nevertheless, GFPNbCy3Cy5 proved effective at binding GFP-Cnp1 in *S. pombe* cells and could be made to blink using a standard *d*STORM oxygen scavenging buffer. This demonstrated the construct's potential use a tool for SMLM, especially since many proteins are GFP tagged and could be imaged using this construct relatively simply compared to generating Nbs to specific proteins of interest, as previously reported (Pleiner et al. 2015). Characterisation of the multicolour and quantitative potential of this construct was next needed

4 Photophysical characterisation of a doubly labelled anti-GFP-Nb for STORM imaging

4.1 Introduction

Chapter 3 demonstrated that an anti-GFP-Nb (GFPNb) identified by Kubala et al (2010) could be successfully labelled with a linking peptide, covalently connected to either a Cy3 and Cy5 fluorophore pair or an Alexa488 and Cy5 fluorophore pair. The two dyes are close in space so we hypothesised that they would be able to act as an activator-emitter pair for STORM experiments as previously reported (Huang, Jones, et al. 2008; Bates et al. 2012; Bates, Blosser, and Zhuang 2005; Bates et al. 2007). However, this needed to be tested as part of a thorough characterisation of these reagents as linking peptides have not previous been used in STORM constructs. This characterisation needed to include testing of whether or not the 1:1:1 nature of the GFPNb-activator-emitter will prove to be a superior probe for dSTORM and STORM experiments.

In this chapter, we determined whether the GFPNbCy3Cy5 STORM construct is suitable for STORM imaging by measuring the activation power dependence and crosstalk parameters, using conventional randomly labelled Abs as controls. We then characterised the photophysics of the construct to assess its suitability for quantitative imaging.

4.1.1 STORM image acquisition

STORM data was collected using a continuous read-out 641 nm laser (in the case of Cy5) constantly illuminates the sample and images are collected at a frequency of 20-100 Hz. A single frame pulse of activator light is applied every 3-10 frames, which causes an increase in the number of localisations immediately afterwards. This is continued until all the fluorophores are bleached or sufficient localisations obtained. A scheme of the acquisition is shown in **Figure 4.1A** and the respective video collection is depicted in **Figure 4.1B**.

As previously described in **section 1.3.3**, STORM uses activator dyes to increase the on rate of cyanine emitter dyes that already blink; Cy5, Alexa647, Cy5.5, Cy7 and Cy7.5. This means that emitters in a sample blink stochastically at a certain rate, but when an appropriate pulse of laser light is shone on the activator it brings more of the emitters out of the dark state than would be expected at random.

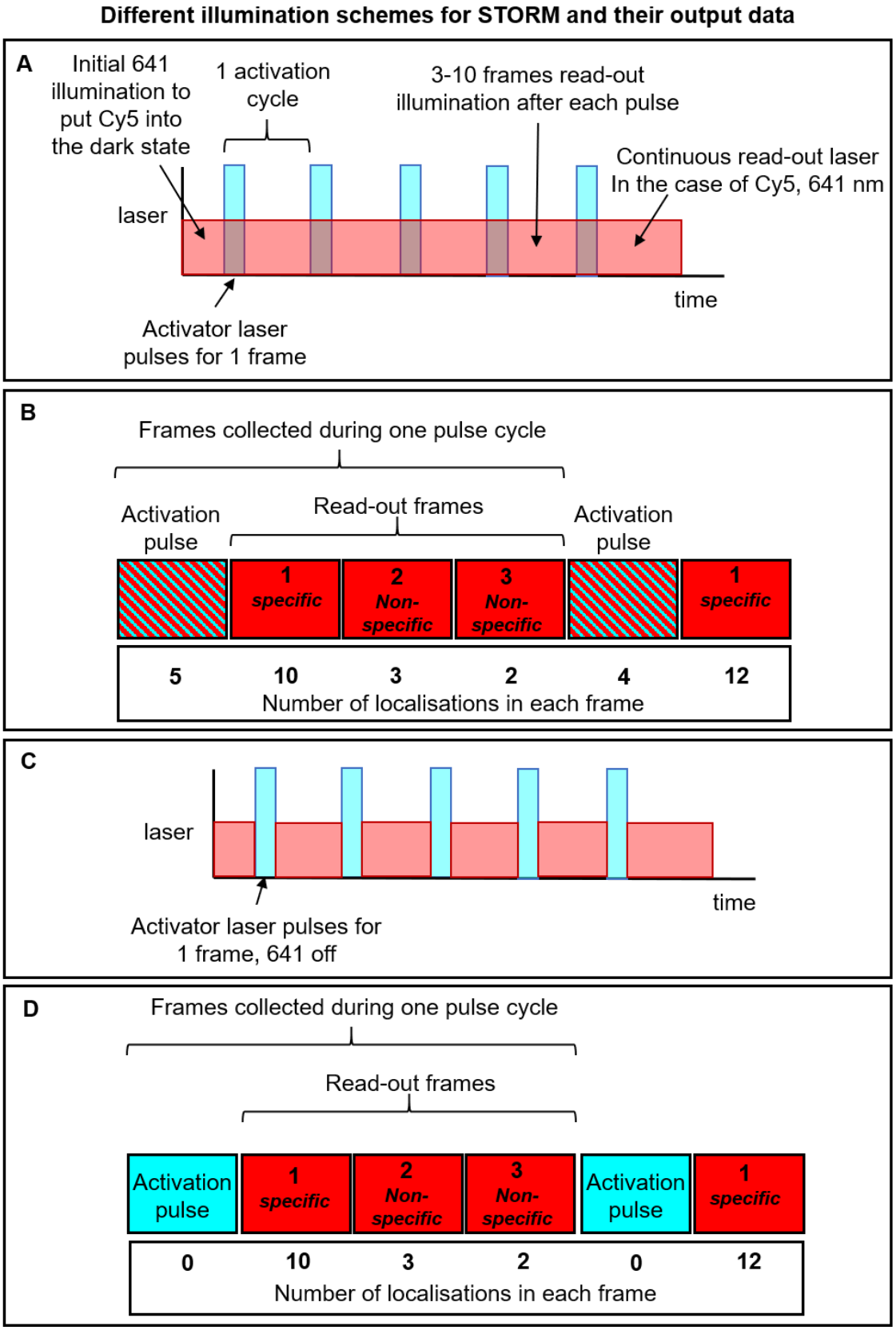


Figure 4.1 - Schemes for one-colour STORM imaging and data collection. ***A*** *Scheme for collection with continuous read-out laser illumination* ***B*** *One cycle of frames collected using scheme* ***A****, with three observation frames in between activation pulses, and pulse frames filled with pattern to indicate 2 lasers on simultaneously.* ***C*** *Acquisition scheme used on Nikon N-STORM for data collection where read-out laser is turned off during activation laser pulses.* ***D*** *A cycle of frames collected during acquisition scheme* ***B.***

4.1.2 STORM data analysis

4.1.2.1 SMLM algorithms from RapidSTORM and PeakFit

STORM data analysis is usually done by specifically written software such as Insight3 and that which is pre-installed on Nikon N-STORM instruments. But ImageJ plugin PeakFit (Herbert 2014) was adapted by its author Dr Alex Herbert (University of Sussex) to allow STORM analysis. It was first necessary to understand the algorithms involved in picking spots which are similar for all SM imaging.

RapidSTORM and PeakFit work by identifying diffraction limited PSFs computing super-resolved co-ordinates from the videos. Details of SR algorithms can be found in the original publications (Herbert 2014; Wolter et al. 2010, 2012). Briefly, both algorithms first reduce noise with a smoothing mask of 5x5 pixels (at least 3x3 for PeakFit) and select candidate maxima using non-maximum suppression, which looks for pixels that are brighter than all those surrounding it. A Gaussian is then fitted to each candidate maximum using methods based on the Levenberg-Marquardt algorithm (Marquardt 1963). If after ~3 iterations fitting is unsuccessful the maximum is discarded. The Fitted Gaussians are then filtered by different criteria. RapidSTORM filters using a minimum size (3 or more pixels across) and an amplitude threshold - a number of counts per 5x5 pixel square (Wolter et al. 2010). PeakFit does this slightly differently and filters based upon signal-to-noise ratio (SNR; calculated by measuring the area under the fitted Gaussian and subtracting the calculated noise per frame), precision, Gaussian width, minimum number of photons and coordinate shift. The two most important threshold values for the work here were the amplitude threshold for RapidSTORM and the SNR for PeakFit. These threshold values must be carefully chosen, because low thresholds pick up background and noise, which negatively skews data, and high thresholds reduce detection efficiency of real localisations. The optimal threshold varies between microscope setups, different laser powers, imaging modes, dyes and between different days (for example subtly different laser alignments) so it was important to always asses these parameters carefully.

The list of X and Y co-ordinates generated by SR software can be used to reconstruct a dSTORM or PALM image. STORM analysis, however, only uses the localisations that appear immediately after an activation laser pulse, localisations in other frames are deemed non-specific, as shown in **Figure 4.1B** and **D.** This is because emitter dyes like Cy5 blink stochastically throughout the acquisition video, with a roughly constant number of localisations per frame and the activation pulse increases its on-

rate, so that a greater number of localisations are seen immediately after pulse activation. For single colour STORM images (containing a single activator-emitter pair e.g. Cy3Cy5) the localisations deemed specific can be plotted in an SR image.

4.1.2.2 Crosstalk analysis for multicolour STORM imaging

For multicolour STORM, analysis is more complex because the same emitters must be assigned a channel based on timing of the respective activation pulse. Dual-colour acquisition uses two differently coloured activation pulses cycled alternately, as shown in **Figure 4.2**.

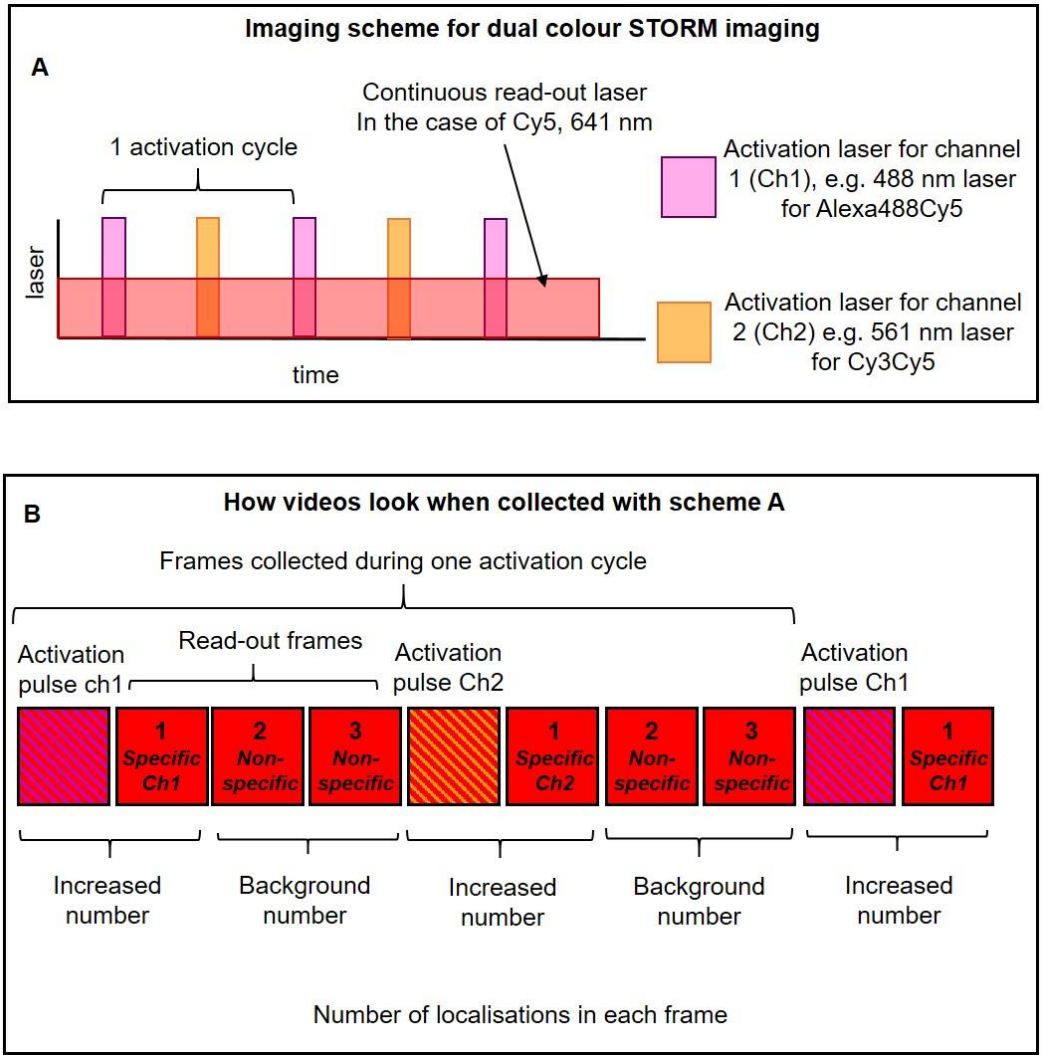


Figure 4.2 – Scheme for dual colour STORM imaging and data collection for a mixed sample containing Alexa405Cy5 and Cy3Cy5. A *Scheme for dual activation data collection with continuous read-out laser illumination* ***B*** *One cycle of frames collected using scheme* ***A****, with three observation frames in between activation pulses of different colours, an indication of the number of localisations per frame is provided underneath.*

Consider imaging the activator-emitter pairs Alexa405Cy5 and Cy3Cy5, activated specifically by 405 nm and 561 nm lasers, respectively. Specific activation of Cy3Cy5 by 561 nm is relative to stochastic blinking and non-specific activation caused by a 405

nm pulse, and vice versa. Also, 405 nm and 488 nm lasers used to excite the Alexa405 and/ and Alexa488 activators cause direct excitation of Cy5 excitation (Heilemann et al. 2005, 2008). These factors mean that the probability of specific activation is never 100%. In previous STORM work, specificity was measured to be >80% in most cases (Bates et al, 2008, 2012). Assignment of localisations to the correct channels was done using an algorithm that considers: the laser pulse identity, the number of nearby localisations observed in each channel (usually ~35 nm radius, (r)) and a measured channel crosstalk parameter. This algorithm uses the apparent density of two species (1 and 2), and reasons that each is equal to the true density plus a few incorrectly assigned molecules, which is a function of the incorrect molecules' density, and characterised by the crosstalk of that channel, as stated below.

$$D_1(r) = d_1(r) + C_{21}d_2(r)$$

Equation 4.1, (Bates et al. 2007, 2012)

$$D_2(r) = d_2(r) + C_{12}d_1(r)$$

Equation 4.2**,** (Bates et al. 2007, 2012)

D_1 and D_2 are apparent densities of STORM pairs 1 and 2, respectively, d_1 and d_2 are the corresponding true densities of each species and C_{12} and C_{21} are the related crosstalk values. As reasoned by Bates et al (2007, 2012), the probability that a localisation is correctly assigned to channel 1 ($P_{correct,\ ch1}$) is therefore related to the apparent and true densities of species 1 as in **Equation 5**.

$$P_{correct,\ ch1} = \frac{d_1(r)}{D_1(r)}$$

Equation 4.3, (Bates et al. 2012)

Substituting **Equation 4.3** with **Equation 4.1** and **Equation 4.2**, produces an expression in terms of only the crosstalk values and the measured apparent densities of species 1 and 2, each of which can be measured, so $P_{correct,Ch1}$ can be solved.

$$P_{correct,\ ch1} = \left(\frac{1}{1 - C_{12}C_{21}}\right) * \left(1 - \frac{C_{21}D_2(r)}{D_1(r)}\right)$$

Equation 4.4, (Bates et al. 2012)

Thus, the dots are assigned probabilistically; on the likelihood that they contain a specific activator and how specifically that activator-emitter pair is activated by each

laser. It is therefore essential to measure the crosstalk values of each STORM pair. Which is done using a rearrangement of **Equation 4.2**.

$$C_{12} = \frac{D_2(r) - d_2(r)}{d_1(r)}$$

Equation 4.5

In a sample containing only one STORM pair, for example pair 1, d_2 is equal to zero and so **Equation 4.5** becomes **Equation 4.6**.

$$C_{12} = D_2/d_1$$

Equation 4.6

C_{12} can be calculated using the number of new localisations activated by the channel 1 laser (d_1) and the number of new localisations activated by channel 2 (D_2), which is not specific to pair 1. The equivalent can be done for pair 2 using **Equation 4.1**. This crosstalk parameter is also a good measure of how specifically a STORM pair is activated by its respective laser – the more specific pair 1 is, the lower the D_2 value becomes so C_{12} is small, although it is unlikely to reach zero because spontaneous Cy5 blinking is always likely to occur and will be attributed to the non-specific laser at the same rate as background.

In published work, software developed in the Zhuang group – Insight3, has been used for STORM analysis, but this is not openly available, hence the adaptation of PeakFit. For some work RapidSTORM was used for analysis that did not require channel assignment, because we have more experience in using it, and crosstalk calculations and channels assignment were carried out with PeakFit.

4.2 Results

4.2.1 SM surface development

To obtain SM data that could be analysed for quantitative photophysics, SM surfaces are used. They are glass slides coated in various layers to allow attachment of fluorophores to the surface, so that fluorophores are spatially separated and can be imaged with total internal reflection (TIRF), which has lower background fluorescence from above the focal plane than highly inclined and laminated optical sheet illumination (HILO, sometimes referred to as near-TIRF), which is used for imaging cell sections more than ~200nm above the coverslip. It was important to optimise the SM surfaces for data collection, to minimise noise and obtain good fluorophore density.

Originally, STORM probes were characterised using DNA oligos affixed to the coverslip using biotin, as shown in **Figure 4.3A** (Heilemann et al. 2005; Bates, Blosser, and Zhuang 2005). However, since our protein construct could not be attached in this way, we investigated several other approaches. An alternative approach was to use dried poly-L-lysine, to stick the protein constructs directly to, or so stick the GFP or equivalent epitope to the poly-L-lysine surface and coat it with BSA as shown in **Figure 4.3B** and **C**. Both SM surface designs were tested with GFPNbCy3Cy5, GFPNbCy3Cy5 and randomly labelled fluorescent Abs.

The STORM constructs stuck to the surface in both designs, but the BSA and epitope coated surface allowed better control of the fluorophore density on the surface. This was because poly-L-lysine is heavily cationic and proteins stick well to it, in a way that was hard to control, whereas, use of BSA limited the extent of STORM constructs sticking to the surface so more reliable fluorophore densities were obtained. Additionally, the BSA method prevented the fluorophores interacting directly with the poly-L-lysine, which may have affected their photophysics through charge interactions which are important for fluorescence in cyanines. This type of surface was always used subsequently.

Surface labelling density is always important in SM characterisation experiments because all experiments have background localisations that skew data and require fluorophores to be sufficiently spread out to ensure the PSFs do not overlap (i.e. are not within the diffraction limited area). The method of using BSA and a suitable epitope was most effective for controlling the fluorophore density in our case.

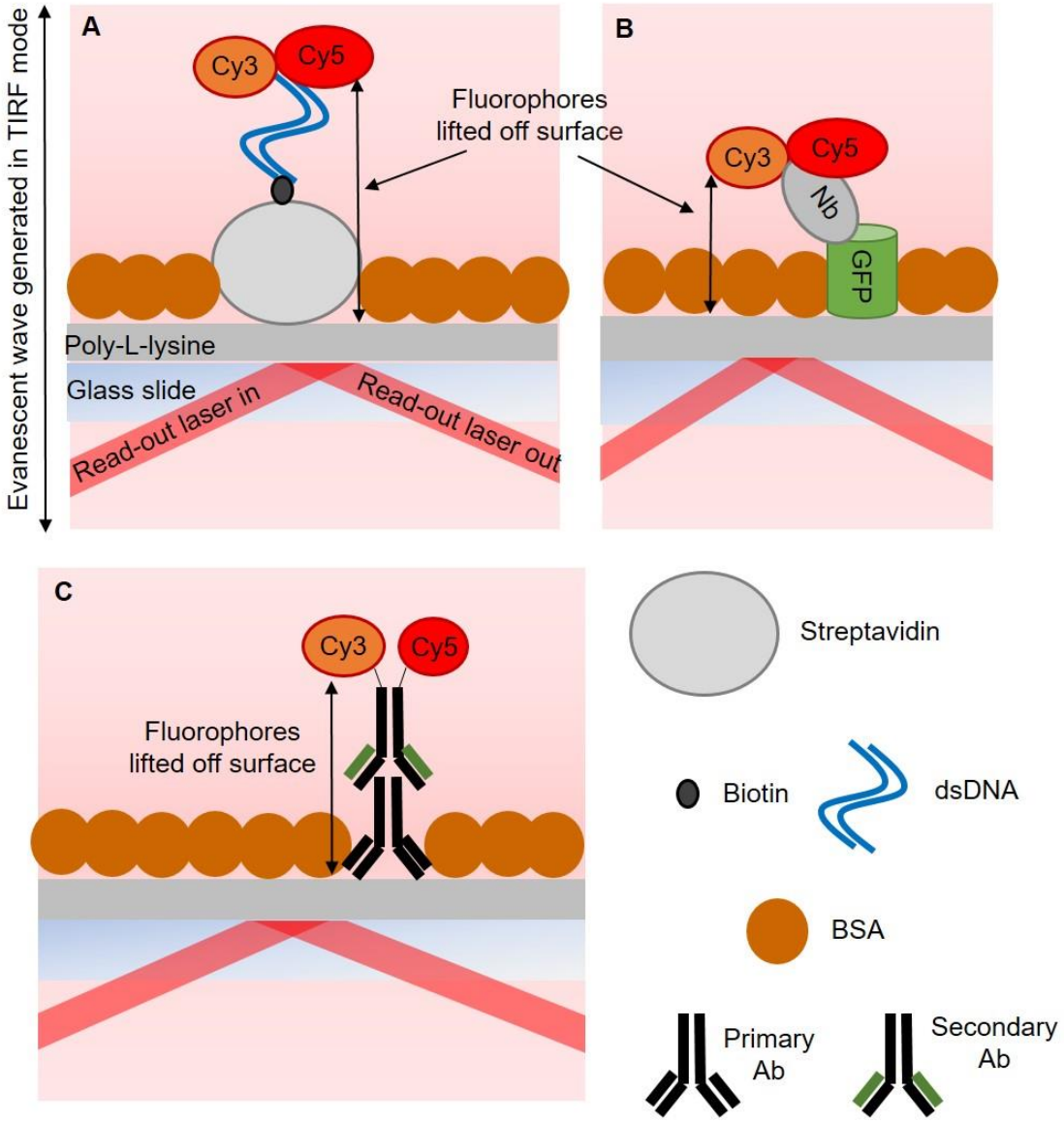


Figure 4.3 - Different schemes for SM surfaces for TIRF imaging. A *Biotinylated DNA labelled with fluorophores and streptavidin.* ***B*** *A GFP surface for use with GFPNbCy3Cy5.* ***C*** *An Ab surface to attach fluorescently labelled secondary Abs to.*

4.2.2 Validation SM analysis thresholds for RapidSTORM and PeakFit Setting

For all SR imaging, it is critical to set a reliable threshold for picking fluorescent spots from raw videos, to do this it is important to understand how the spot picking algorithms work, which is described in **section 4.1.2.1**. There are many ways to optimise a threshold to raw data, for example using the raw integrated density of dots compared to background as a starting point and then to tune the exact value by counting the number of localisations in a small set of data and comparing that to the output by fitting programs such as RapidSTORM or PeakFit.

Here a suitable RapidSTORM amplitude threshold was initially estimated by measuring the difference between integrated density between a small region of interest (ROI) that contained a spot and one that did not, in a GFPNbCy3Cy5 sample because each dot should only represent one Cy5 fluorophore. Several of these were measured and showed a difference of ~100 to 300 photons. SM raw maximum intensity traces and fitted intensity traces localised with an amplitude threshold of 25,000 counts were plotted against frame number (an example is shown in **Figure 4.4A** and **B**). All maxima for this spot with a maximum pixel intensity of more than ~4000 counts (35 photons) have been localised by RapidSTORM with a global amplitude threshold of 25,000. This indicates that 25,000 is a good threshold for these data. That threshold was then applied to all the GFPNbCy3Cy5 and AbCy3Cy5 data, to allow fair comparison. **Figure 4.4B** also gives an indication as to the mean fitted intensity due to a single Cy5 molecule, which appears to be approximately 60,000-80,000 counts (~500-750 photons) – the approximate average height of the peaks. This is important in characterising physical properties as will be discussed later.

Additional details of the imaging set up and analytical parameters are described in **section 6.7**.

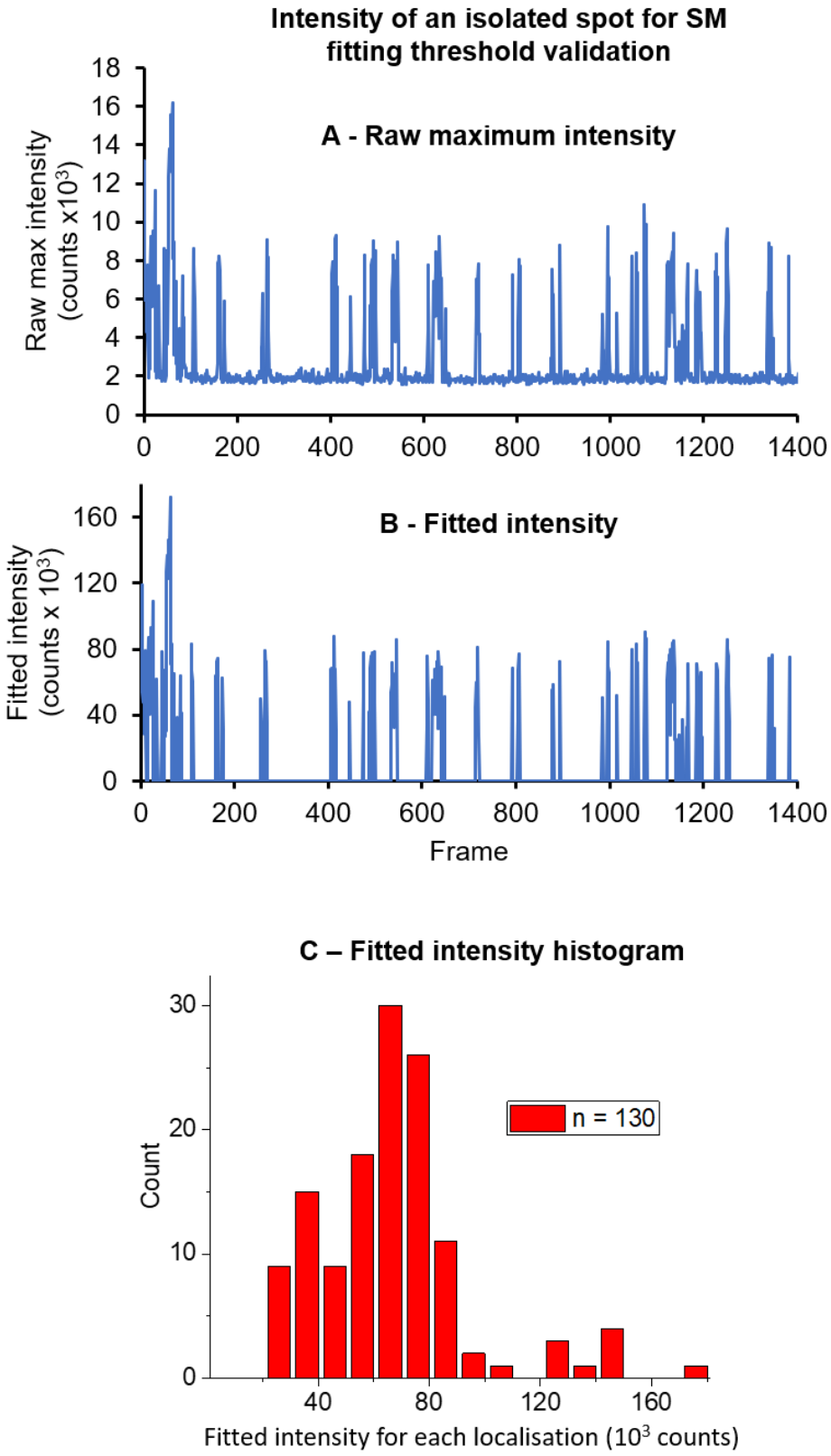


Figure 4.4 – Validation of RapidSTORM global amplitude threshold. ***A*** *Raw maximum intensity in the raw video file shows where potential localisations are.* ***B*** *Fitted intensity from RapidSTORM using a global threshold of 25,000.* ***C*** *Histogram of fitted intensities for each localisation. The bin width was calculated to be close to optimal for the data* (Izenman 1991).

Since the crosstalk algorithm and dual colour STORM imaging are PeakFit plugins, it was also necessary to set a threshold in PeakFit that as closely as possible produced

the same localisation lists to RapidSTORM at the threshold we set – 25,000. PeakFit uses different filtering parameters to RapidSTORM, the most important of which is the signal-to-noise ratio (SNR). To calibrate this to the global RapidSTORM amplitude threshold of 25,000 the PeakFit parameter was varied for several videos and the output number of localisations was initially compared. A SNR value of 30 had the closest number of localisations to RapidSTORM, within 2% on average between different videos. An example region of interest (ROI) was fitted using these threshold parameters for RapidSTORM and PeakFit for comparison, and are shown in **Figure 4.5**. The two images showed a high close resemblance, with all the brightest areas, representing many localisations, in common between the two, although some dimmer localisations in the RapidSTORM image are apparently not present in the PeakFit one.

We also quantified the number of false positives – the number of localisations detected on a SM surface that should not contain any fluorophores, in this case poly-L-lysine, BSA and epitope such as PA-GFP or primary Ab. Ideally, there would be no false positives above noise, but this is rarely the case in practice. **Figure 4.6A** shows how many localisations were found in one field of view on a blank slide containing the PA-GFP at different SNR thresholds. The threshold identified previously was SNR = 30, at this value there were 38 localisations from 13 'fluorophores'. This compared to 38919 localisations for a slide containing GFPNbCy3Cy5, indicating approximately 1000-fold more localisations in the experimental sample. Two representative fields of view in **Figure 4.6B** and **C** for the blank slide and this experimental sample also demonstrate how the density of molecules is far greater in the experimental sample. This is strong evidence that whilst there are some false positives as a result of the surface itself, the vast majority of detected localisations are due to the fluorescent constructs, so the false positives should have negligible effect on the overall statistics of the experimental samples at this density.

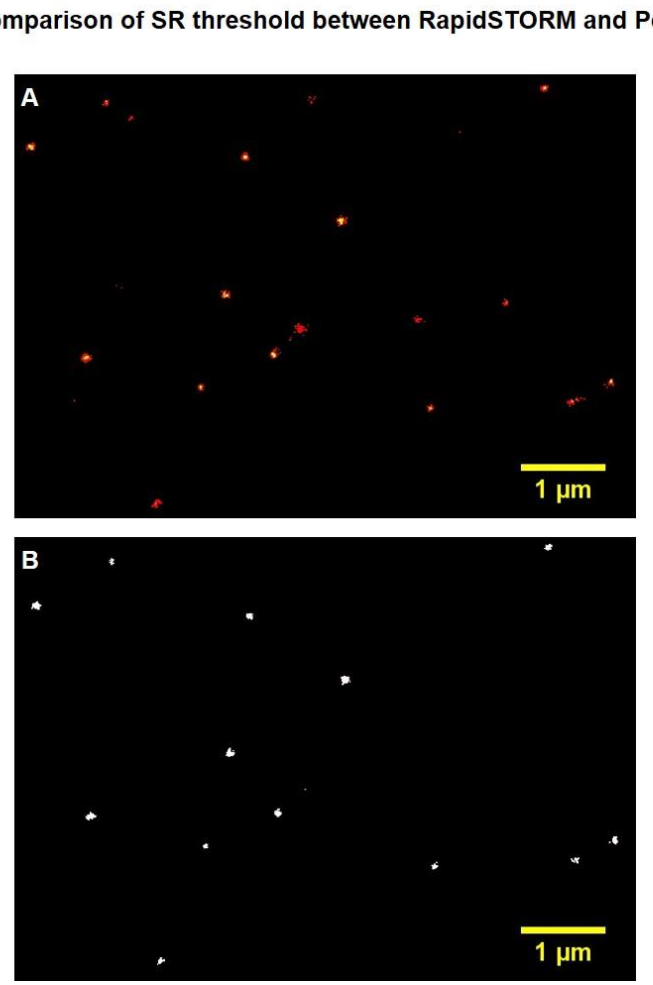


Figure 4.5 – Validation of PeakFit SNR threshold. *Two SR ROIs from images where the same raw dataset was fitted using* ***A*** *RapidSTORM and* ***B*** *PeakFit. Both images were plotted with 10 nm pixel width so that they are comparable, the scalebar represents 1 µm.*

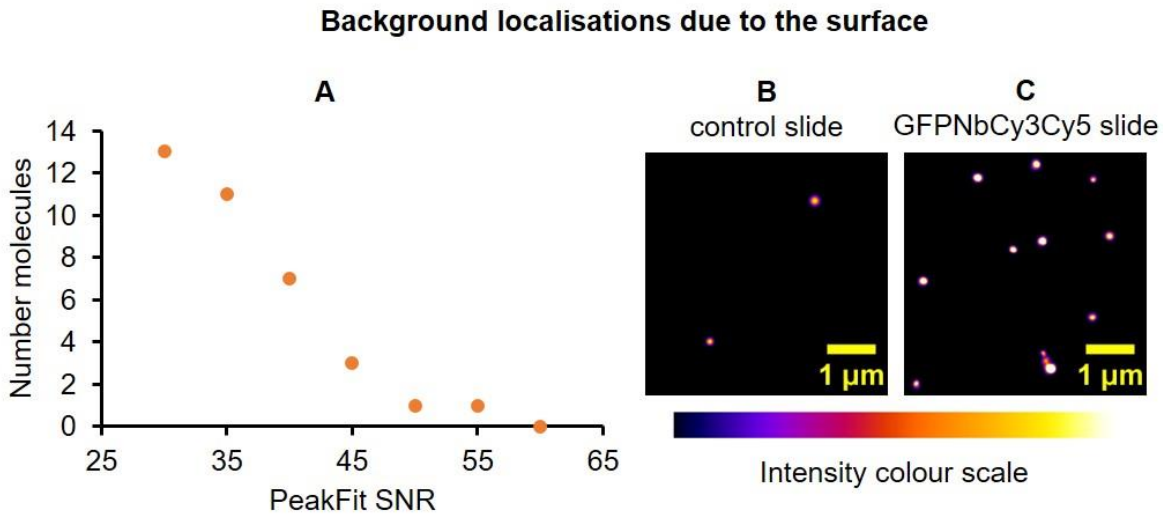


Figure 4.6 – Localisations produced on blank surfaces. *A Plot of the number of molecules fitted to a dataset from a blank SM surface using PeakFit, with different thresholds. B and C representative 5 µm x 5 µm areas of SR images of B a blank control slide and C a slide with the GFPNbCy3Cy5 fluorophores added at suitable imaging density. The scalebar indicated is 1 µm and the colour scheme for intensity is shown beneath the images*

4.2.3 GFPNbCy3Cy5 is a functional STORM construct

Having demonstrated blinking of Cy5 in GLOX oxygen scavenging system, it was necessary to verify whether the Cy5 blinking rate was dependent upon the activator dye in our constructs since this is essential in STORM imaging. Previously, this has been done by measuring the k_{on} due to the activation laser (Bates, Blosser, and Zhuang 2005). In these experiments samples were illuminated with 641 nm laser light to turn Cy5 to the dark state, and then the activator laser was switched on for an unspecified time (probably tens of frames). This caused the Cy5 fluorophores to return to the fluorescent state, and to reach a new equilibrium between the dark state and the fluorescent state, driven by the opposing 641 nm laser and the activator laser. k_{on} was calculated from the number of Cy5 molecules fluorescing at equilibrium. We attempted to replicate this method, but were unable to reach an equilibrium state unless the SM surface was completely covered with fluorophores, which we judged as incomparable to imaging spatially separated fluorophores. Furthermore, this method would have to be done on separate fields of view for separate activation lasers due to likely bleaching of the activator dye under constant illumination. We chose to use a novel method to observe Cy5 activation by the correct laser, which used single activator pulses and 29 subsequent observation frames (see **Figure 4.7A**), because it allowed us to use conventional SM surfaces with spread out fluorophores to collect data for quantitative characterisation; and use multiple activation lasers in the same video to assess specific activation, non-specific activation and crosstalk.

Theoretically, the number of fluorophores activated by a pulse is proportional to the activation laser power. Therefore, it was decided to use the increase in number of localisations after a pulse as an indicator of Cy5 activation by the activator laser. We related the number of fluorophores visible after activation (N_{max}) to the number on before that pulse (N_{eq}) using **Equation 4.7**, a value we called the 'fractional increase' (FI).

$$FI = \frac{N_{max}}{N_{eq}}$$

Equation 4.7

When FI is approximately 1, N_{max} is not larger than N_{eq} so there is no activation by that laser. For a specific laser at a suitable power to cause activation, N_{max} was expected

to be several times larger than N_{eq} thus FI would be larger than 1, perhaps up to 10 for an obvious activation.

Alternative formulae representing degree of activation could have been used and include other control variables such as density of fluorophores, and could relate to fraction of molecules activated, for example, **Equation 4.8**.

$$Fraction\ of\ molceules\ activaed = \frac{N_{max} - N_{eq}}{N_{tot}}$$

Equation 4.8

where N_{tot} is the total number of fluorophores in the field of view. Whilst this sort of value may be more meaningful, N_{tot} has complications associated with it such as activator and emitter bleaching which change the number of decrease the number of fluorescent Cy5 molecules throughout a video. Future work could consider these variables, but FI was sufficient for demonstrating STORM pair activation specificity here.

SM surfaces were prepared and imaged using constant read-out laser illumination and activation pulses every 30 frames, as illustrated in **Figure 4.7A**. Analysis was then carried out by fitting PSFs to dots in the video with RapidSTORM and taking the modulus of the frame number. Plotting number of localisations by the modulus of the frame number allows overlaying of multiple activations cycles, relative to the pulse activation frame. The total number of localisations in any frame relative to the activation pulse, can then be counted, as shown in **Figure 4.7B**. An example set of results with important measurement points is shown in **Figure 4.7B**, where the fractional increase due to 561 nm, 488 nm and 405 nm were found to be 6.2, 2.5 and 1 respectively. This method allows direct comparison of the activation caused by specific and non-specific activation lasers, no activation is indicated by FI = 1.

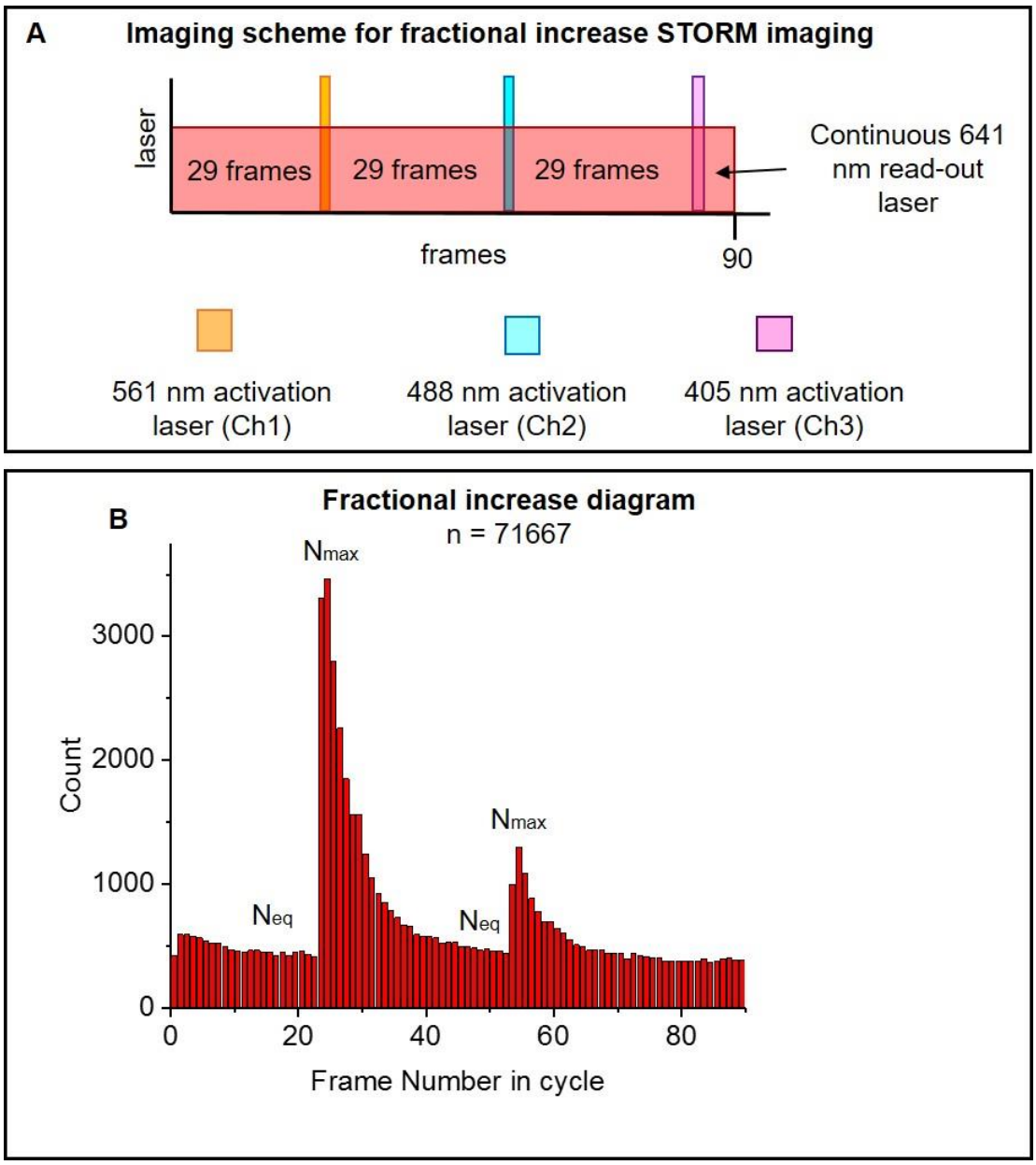


Figure 4.7 – "Fractional increase". A *Laser scheme for data collection.* ***B*** *Histogram of all localisations in a resulting video of AbCy3Cy5 with a pulse cycles of 90 frames. The modulus (90) of the frame number was taken and the localisations binned by which frame they appeared in. The horizontal axis is aligned with A to show when each pulse occurred. The positions of* N_{max} *and* N_{eq}

Four constructs were used to measure the power dependence of this fractional increase: two secondary Abs labelled with Alexa405Cy5 and Cy3Cy5 (average number, 0.88 Cy5 per Ab on average), GFPNbCy3Cy5 and GFPNbAlexa488Cy5, described in **Chapter 3**. The Abs were labelled using NHS esters of Cy5 and the respective activator dyes via solvent facing lysine resides as described in **section 6.6.5** and previous STORM work (Ricci et al. 2015; Lakadamyali et al. 2012). They were used as positive controls because they should be fully functional STORM pairs. It was expected that the fractional increase of Cy5 localisations would increase with respective activation laser power in all three cases if the STORM pairs are functional.

The Cy3Cy5 constructs should only show this pattern for the 561 nm laser, the AbAlexa405Cy5 construct should only show that pattern for the 405 nm laser and the Alexa488Cy5 construct should only respond to the 488 nm laser. Non-specific lasers were expected to have little to no effect on fractional increase.

The fractional increase power dependence for AbAlexa405Cy5 in response to 405 nm and 561 nm lasers is shown in **Figure 4.8**. It shows that at low power (~0.1-0.8 W/cm^2) 405 and 561 nm pulses there is very little fractional increase – there are not many more localisations than with just Cy5 illumination, however, in the range ~0.8-4 W/cm^2 there is considerable fractional increase for 405 nm activation, but no increase with the 561 nm pulses. The variation between repeats is also very low. This is good evidence that the Ab405Cy5 pair is activated by 405 nm, but not 561 nm light – so the STORM pair is functional and specific. This also indicates that the fractional increase is a valid measure of responsiveness to activation.

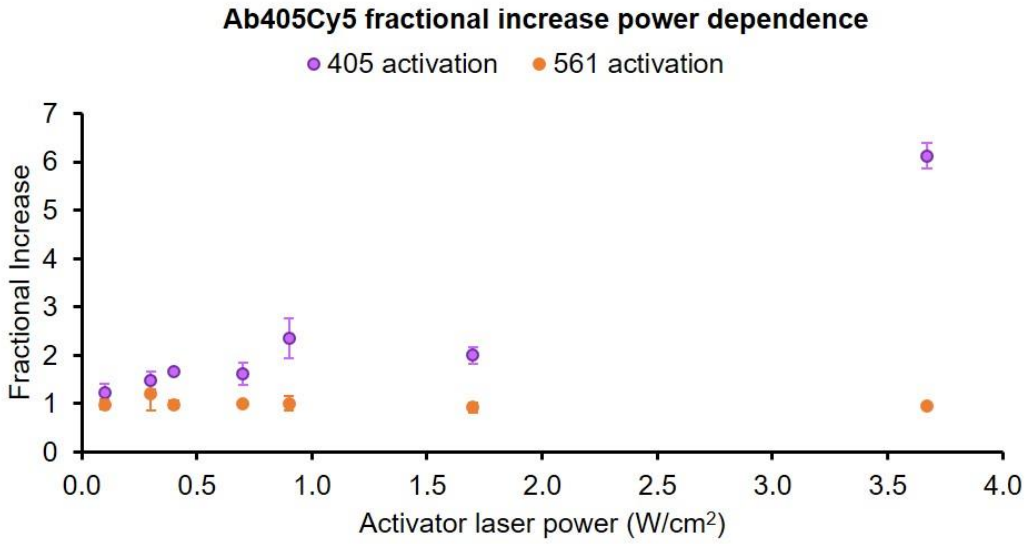


Figure 4.8 – Fractional increase power dependence for Ab405Cy5 positive control.

Fractional increase power dependences were also carried out for the AbCy3Cy5 and GFPNbCy3Cy5 constructs using three activator lasers instead of two; 405 nm, 488 nm and 561 nm, all of which have been used for STORM pairs as discussed in **section 1.4.2.4**. Different ranges of activator laser powers were used based on preliminary data, which suggested 488 nm light increased FI in the range 1-5 W/cm^2, whereas, 405 nm light only increased FI at higher powers, much higher than we intended to use. Therefore, the 488 nm laser was used in a power rage across spread equally across 1-5 W/cm^2 range and the 405 nm laser was used over a power lower range ~<1.5

W/cm^2, where we expected to image plus one excessively high power measurement. The results are shown in **Figure 4.9**.

In both cases the fractional increase due to 561 nm pulses was clear, with an up to 8-fold increase in the number of localisations per frame in both cases. Both constructs displayed an initial rising fractional increase with 561 nm laser power, but this appeared to plateau above ~0.8 W/cm^2. In the AbCy3Cy5 case the standard deviation increased with power. Both Cy3Cy5 constructs showed little change in fractional increase below 2 W/cm^2 405 nm activation laser power, which indicated that the 405 nm light has no effect of the Cy3Cy5 pair in this power range, although there was an increase at ~7.5 W/cm^2 for the GFPNbCy3Cy5 construct (**Figure 4.9C**), indicating that at higher 405 nm laser power it did have an effect on this construct. Future experiments must use less than 7.5 W/cm^2 to avoid crosstalk.

The 488 nm and 405 nm lasers caused significant activation above ~1 and ~7 W/cm^2 respectively. This was likely due to overlap of the broad Cy3 and Cy5 absorption spectra with these wavelengths, which causes a limited amount of absorption, for example, Cy3 absorbs approximately 9% of 488 nm light, hence it was expected that 488 nm light could cause STORM pair activation some of the time, as seen previously (Bates, Blosser, and Zhuang 2005; Bates et al. 2007). Furthermore, Cy5 in some of its dark states is known to absorb 405 and 488 nm light, which can also return it to the on state (see **Section 1.3.2**; Heilemann et al. 2005, 2008). For these reasons it is important to use as low as possible power 405 and 488 nm lasers in multicolour STORM experiments to avoid crosstalk. In our case it appears using powers ~ 1 W/cm^2, seems sensible to minimise crosstalk with Cy3Cy5.

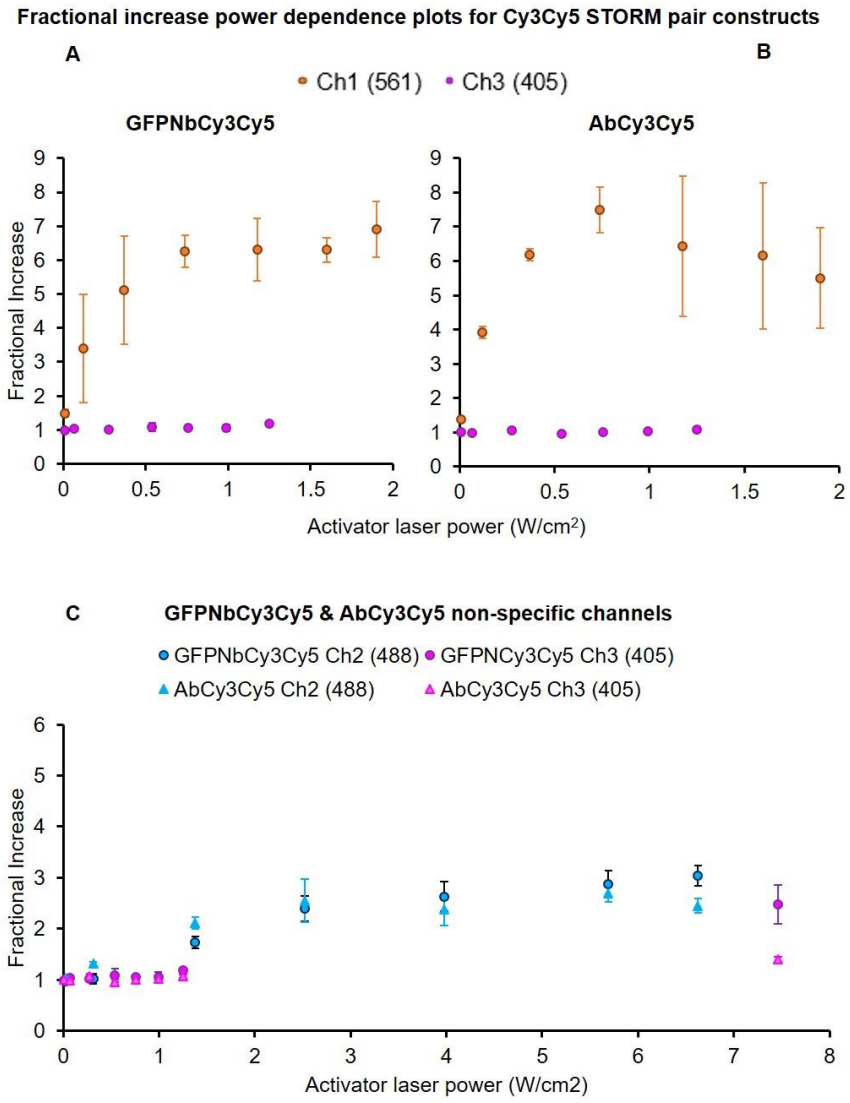


Figure 4.9 – Fractional increase power dependence plots for GFPNbAcy3Cy5 and AbCy3Cy5. ***A*** *and* ***B*** *the power range used with the specific 561 nm and 405 nm lasers.* ***C*** *shows the whole power range for the non-specific activator lasers 488 nm and 405 nm and both constructs. GFPNbCy3Cy5 points are shown with filled circles and AbCy3Cy5 by filled triangles. All error bars are sample standard deviation from 3 repeats. The 488 nm laser caused significant fractional increase in the range 1.3-7 W/cm^2. This suggested that this laser activates the Cy3Cy5 STORM pair more than 405 nm light and that the 488 laser is best not combined with the Cy3Cy5 pair, so it was not used in further analysis.*

Finally, GFPNbAlexa488Cy5 was characterised with 488 nm laser pulses only, as shown in Figure 4.10. No fractional increase was evident at any 488 nm laser power in the range 0.1-4 W/cm^2, producing a flat line. This indicates that the Alexa488Cy5 pair was not responsive to the 488 nm laser, and so was not a functional STORM pair.

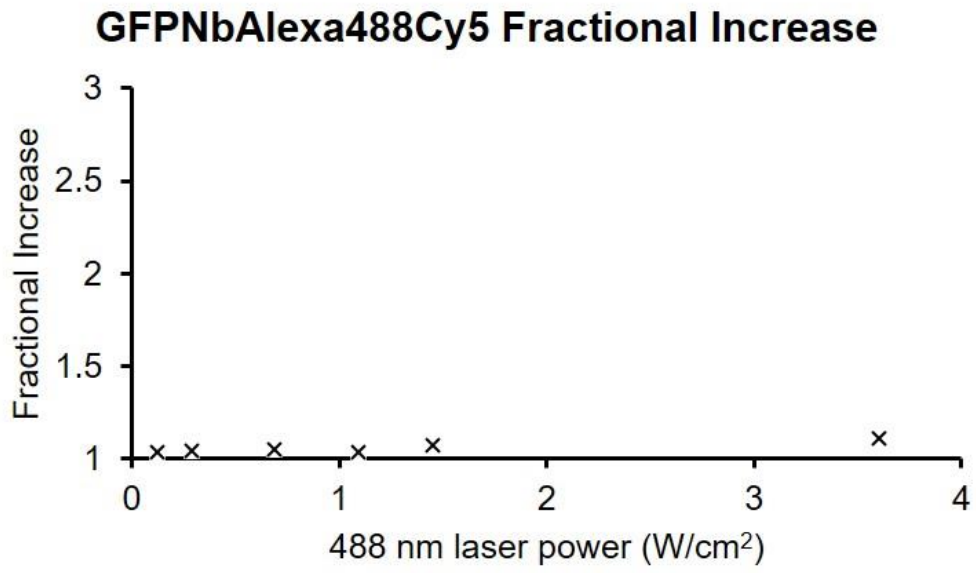


Figure 4.10 ***– Fractional increase power dependence for GFPNbAlexa488Cy5.***

These power dependence experiments strongly indicated that three of the four STORM constructs investigated here were responsive to their respective activation lasers, and appeared to be activation laser specific. This was good evidence those three are functional STORM pairs. GFPNbAlexa488Cy5 did not show any power dependence to 488 nm laser activation, so that pair was judged as not functional. It was not characterised further.

4.2.4 Constructs are specifically activated with low crosstalk

It was important to further investigate the specificity of activation of these three constructs, to verify the indications of the fractional increase experiments. As described in **section 4.1.2.2** and **4.2.3** STORM pair specificity can be measured by k_{on} or by measuring the crosstalk component as if carrying out a dual colour experiment control.

To measure the crosstalk in the same way as previous groups, a PeakFit plugin for ImageJ was written by Dr Alex Herbert (University of Sussex), which used the same algorithm as described by Bates et al (2007, 2012).

The data collected for the previous fractional increase experiments was used as an input for this calculation. The crosstalk algorithm found all tracks that start in the pulse activation frames for each activation laser, in this case 405 and 561 nm, and calculated C_{12} or C_{21} from the measured D_1 and d_2 or D_2 and d_1, using **Equation 4.6**. For simplicity, channel 1 was assigned to 561 nm, channel 2 to 405 nm.

It was expected that a specific STORM pair, for example AbAlexa405Cy5, would have a low C_{21} value indicating low numbers of localisations caused by the 561 nm laser pulse. The data was collected in a power series, increasing both lasers by approximately the same amount for each experiment. This may decrease the crosstalk (increase the specificity) at lower powers, but the non-specific laser would be expected to cause an increase in C_{21} at excessively high powers. Whilst data was collected for the 488 nm laser as well, that data was disregarded from this analysis for simplicity. The results for AbAlexa40Cy5, AbCy3Cy5 and GFPNbCy3Cy5 STORM constructs are shown in **Figure 4.11**.

AbAlexa405Cy5, shows very high crosstalk at low activation laser powers, but after ~0.4 W/cm^2 it decreases as the activator laser power increases. This suggests that the 561 nm laser has little effect on this STORM pair, and that the 405 nm laser only activates specifically above ~0.4 W/cm^2, and is more specific at higher powers. This data shows that at the highest measured power, ~90% of activations of this construct are caused by the 405 nm laser.

The Cy3Cy5 constructs were more specifically activated by 561 nm laser light, as the lowest crosstalk values measured are <0.05, so more than 95% of activations of this pair were caused by the 561 nm laser and less than 5% by the 405 nm laser. The power dependence, however, showed a slight increase in the crosstalk at activator powers >1.5 W/cm^2, which was probably due to the 405 nm laser starting to have an effect in this range as reported in *d*STORM experiments at <1 W/cm^2 (Lampe et al. 2012). In all cases the GFPNbCy3Cy5 had slightly higher crosstalk than the AbCy3Cy5, indicating it was a slightly less good construct for multicolour STORM. This maybe because AbCy3Cy5 has an average of more than one Cy3 per Cy5, so activation was more likely.

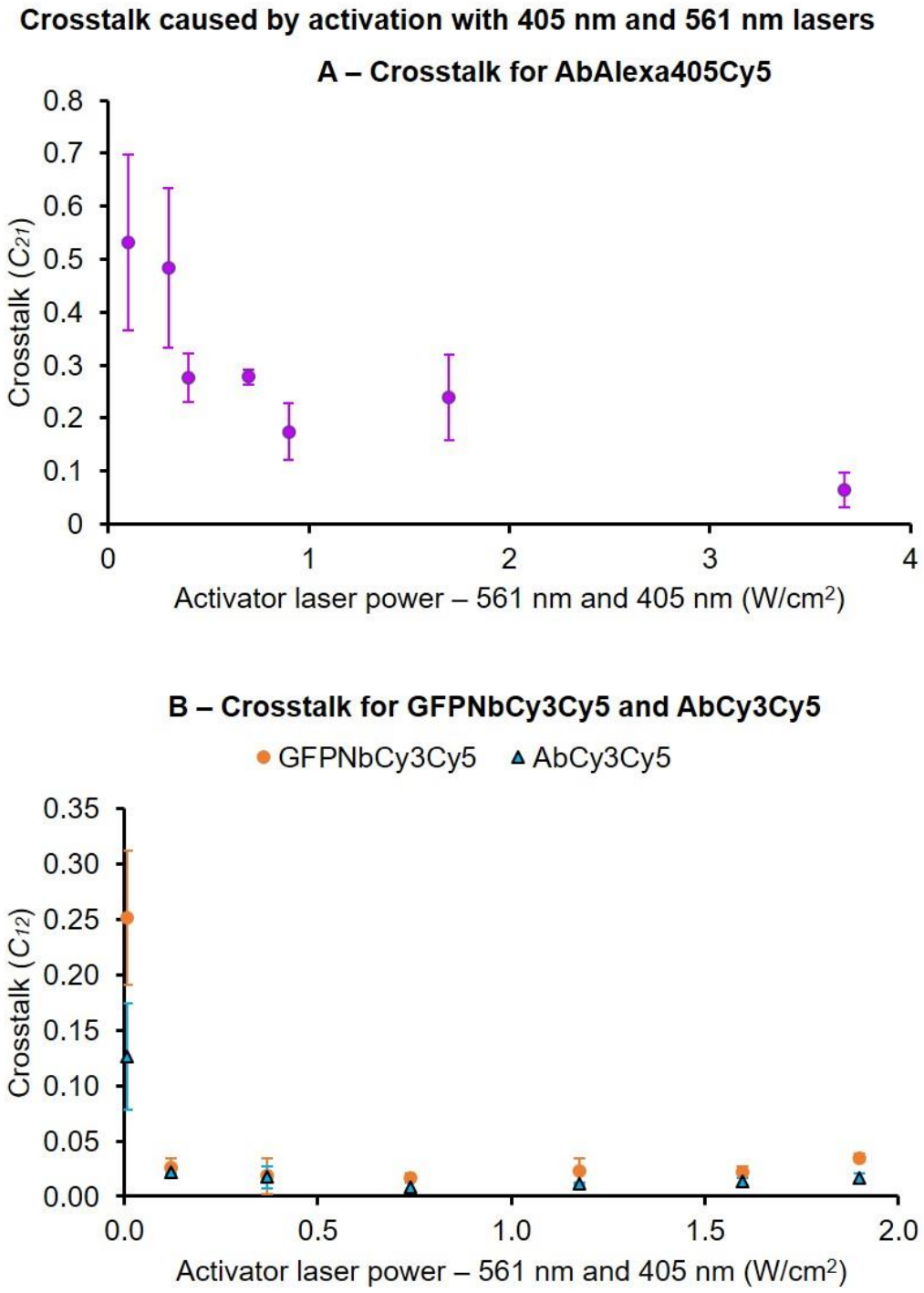


Figure 4.11 – Crosstalk for Cy3Cy5 STORM pair from non-specific laser is generally lower at higher activator laser powers for single STORM pair samples. *Crosstalk measured on single STORM pair samples for* ***A*** *AbAlexa405Cy5, where the 405 nm laser should be specific, the 561 nm laser should be non-specific and* ***B*** *GFPNbCy3Cy5 and AbCy3Cy5 which should be specifically activated by the 561 nm laser, but not the 405 nm laser. Illumination scheme is the same as that used in* ***Figure 4.7****A.*

The crosstalk was also visualised in two colours from these experiments, as shown in **Figure 4.12.** This demonstrated that in a sample of GFPNbCy3Cy5, [crosstalk (C_{12}) = 0.01939] 116 fluorophores are initially resolved, of these 85 are assigned to the correct 561 nm channel and 6 are assigned incorrectly to the 405 nm channel. Although this approximate calculation does not consider all localisations, the effect of crosstalk is

evident. Therefore, even constructs like these with low crosstalk values cannot completely remove false localisations. This will affect both multicolour imaging and quantitative imaging.

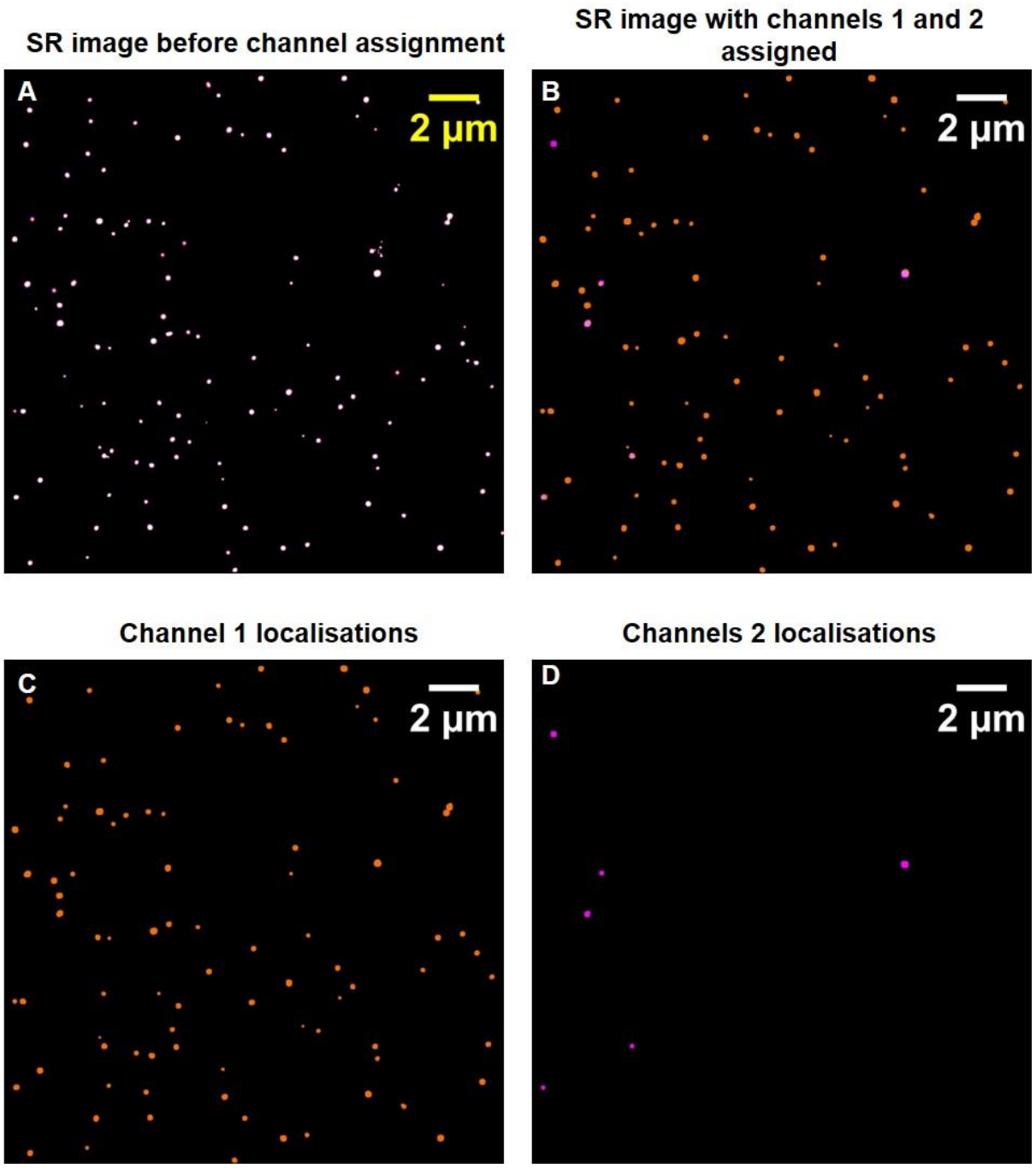


Figure 4.12 – STORM image reconstruction of a sample containing GFPNbCy3Cy5. ***A*** *All SR localisations found using PeakFit.* ***B*** *SR localisations assigned to the 561 nm activation laser (channel 1) and the 405 nm laser (channel 2), using the calculated crosstalk values (C_{12} = 0.01939 and C_{12} = 0), and falsely coloured.* ***C*** *Channel 1 assignments only.* ***D*** *Channel 2 assignments only. All plots have a scalebar of 2 µm, and are plotted at a resolution of 8.6 nm/pixel.*

4.2.5 Single GFPNbCy3Cy5 particles are mostly homogeneous

It is important to understand how individual particles behave and how they relate to average measurements, it was particularly important to determine whether the GFPNbCy3Cy5 when imaged is observed as single homogeneous Cy5 fluorophores. This is important for counting experiments because the more homogeneous detection is, the better calibrations can be done, which can result in greater accuracy and precision. The data collected previously was analysed by studying single particles and measuring the number of photons detected for each localisation. The results for four single particles are shown in **Figure 4.13**, the same measurement of 300 particles is shown in **Figure 4.16A**. The methodology is described in detail in **section 6.7,** briefly, for histograms, single particles were isolated and all maxima localised with RapidSTORM, localisations were traced through time and grouped into blinks. The first and last localisation of each blink was removed and the intensity of each localisation was converted to photons and binned for plotting. For intensity traces, the isolated particle had the maximum intensity for each frame plotted against frame number.

The four representative particles in **Figure 4.13** all have intensity histograms with a main peak and a negative skew. The most common intensity observed is ~600-700 photons, except, particle **B** has a modal intensity of 400-500 photons, hence it is labelled as ‘dim’. This single distribution for all four particles can also be seen in the raw maximum intensity traces, which show peaks of similar height for particles A, C and D, averaging about 7000-9000 counts. These data indicate a homogenous intensity produced by the GFPNbCy3Cy5 construct. Some particles, like **Figure 4.13C** are dimmer than others, which may be due to the environment of that fluorophore being different somehow.

The negative skew in all four particles’ histograms and lower raw maximum intensity peaks indicated that a small number localisations in each trace were relatively dim. It was unlikely that the negative skew was caused by a single Cy5 fluorophore per particle because the maximum intensity traces showed no evidence of stepwise photoactivation or deactivation, as we would expect when two or more fluorophores were present. In some isolated particles, stepwise photoactivation or deactivation was found, as shown in **Figure 4.14A** and **B** but these were in the small minority and were not as convincing as those readily found in the multiply labelled AbCy3Cy5 sample, an example is shown in **Figure 4.14C.** Since the negative skew was unlikely to be caused by a separate number of fluorophore per particle, it was probably the result

photophysical artefacts that could be produced in several ways such as fluorophores flickering on and off or entering the dark triplet state on a timescale faster that detected here.

If one Cy5 produced ~400-700 photons per frame, these data suggest that the vast majority of particles only had one Cy5 because there are very few instances of 800-1400 photons per frame, although A had some localisations that were in this intensity range. This meant that two Cy5 molecules were detected simultaneously in a small minority of cases, and that there probably would be some cases where photoactivation and de-activation were evident. Some examples of both GFPNbCy3Cy5 and AbCy3Cy5 particles are shown in **Figure 4.14**. The insets of raw maximum intensity traces show that the peaks could correspond to one and two fluorophores per particle, but they are very short-lived ~ 5-10 frames (200-400 ms) compared to the AbCy3Cy5 case in **Figure 4.14C** which has a much more obvious photobleaching trace with at least 3 'steps' evident each lasting at least 2s. This stepwise photoactivation and de-activation evidence shows that some cases of two Cy5 fluorophores per particle exist in the GFPNbCy3Cy5 case, but vary rarely were they observed, whereas, in the AbCy3Cy5 case they were common and exaggerated.

It was also possible to use the intensity histograms for 300 GFPNbCy3Cy5 particles combined to quantitatively estimate the degree of labelling particles with more than one Cy5. There were 74 localisations with fitted intensity of 800 photons or more in this dataset, out of 5264 localisations in total, 1.4%. This indicated the extent of double labelling or dimerisation of GFPNbCy3Cy5 is probably ~ 1.4%. So the overwhelming majority is singly labelled.

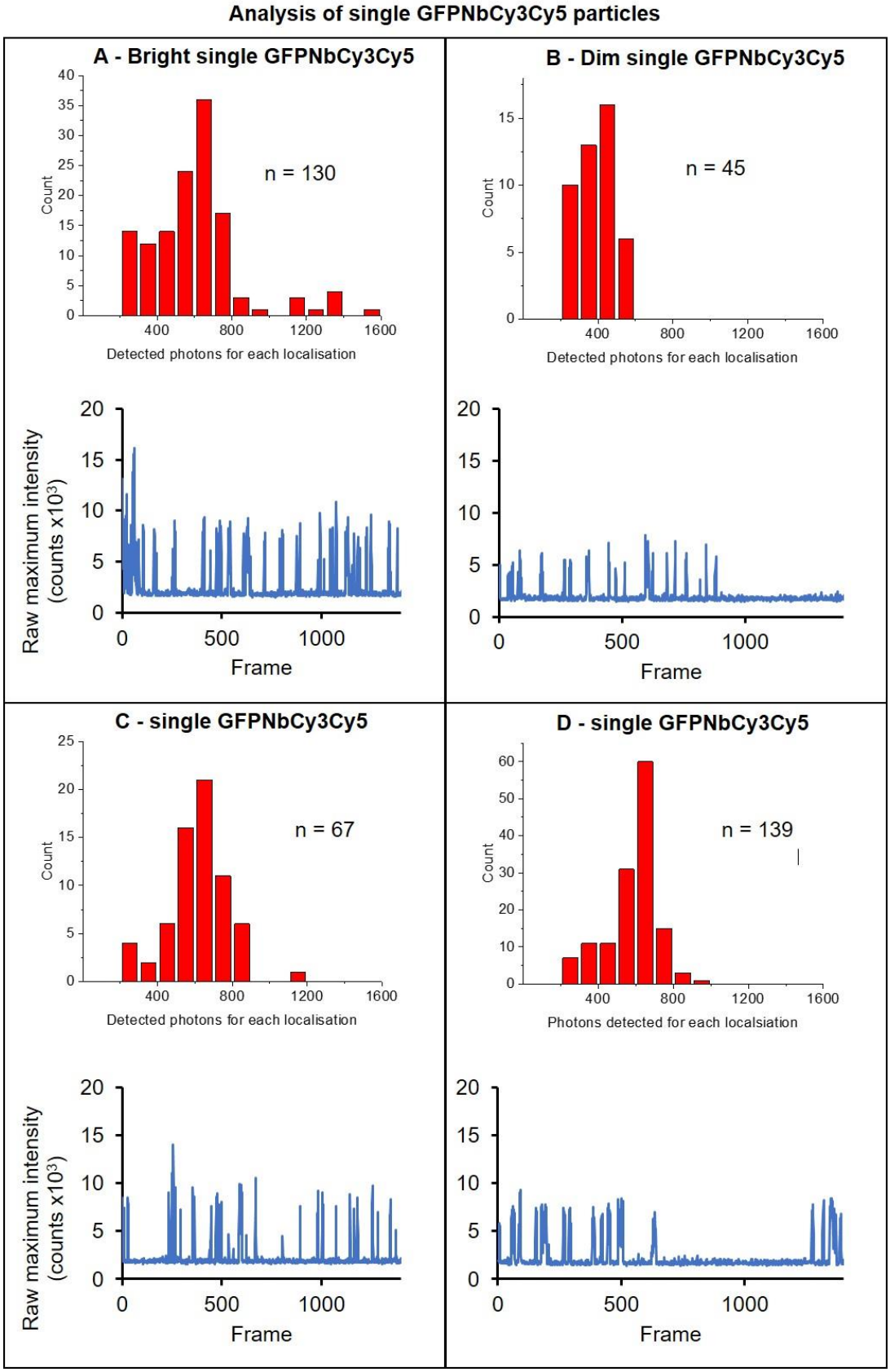


Figure 4.13 – Detected intensity varies between single GFPNbCy3Cy5 particles. Intensity h*istograms and raw maximum intensity traces of four single GFPNbCy3Cy5 particles* ***A****-****D****. All histograms are plotted with the same X-axis scale and all traces are plotted with the same X and Y axes' scales to allow comparison. The number of localisation produced by each particle (n) is shown.*

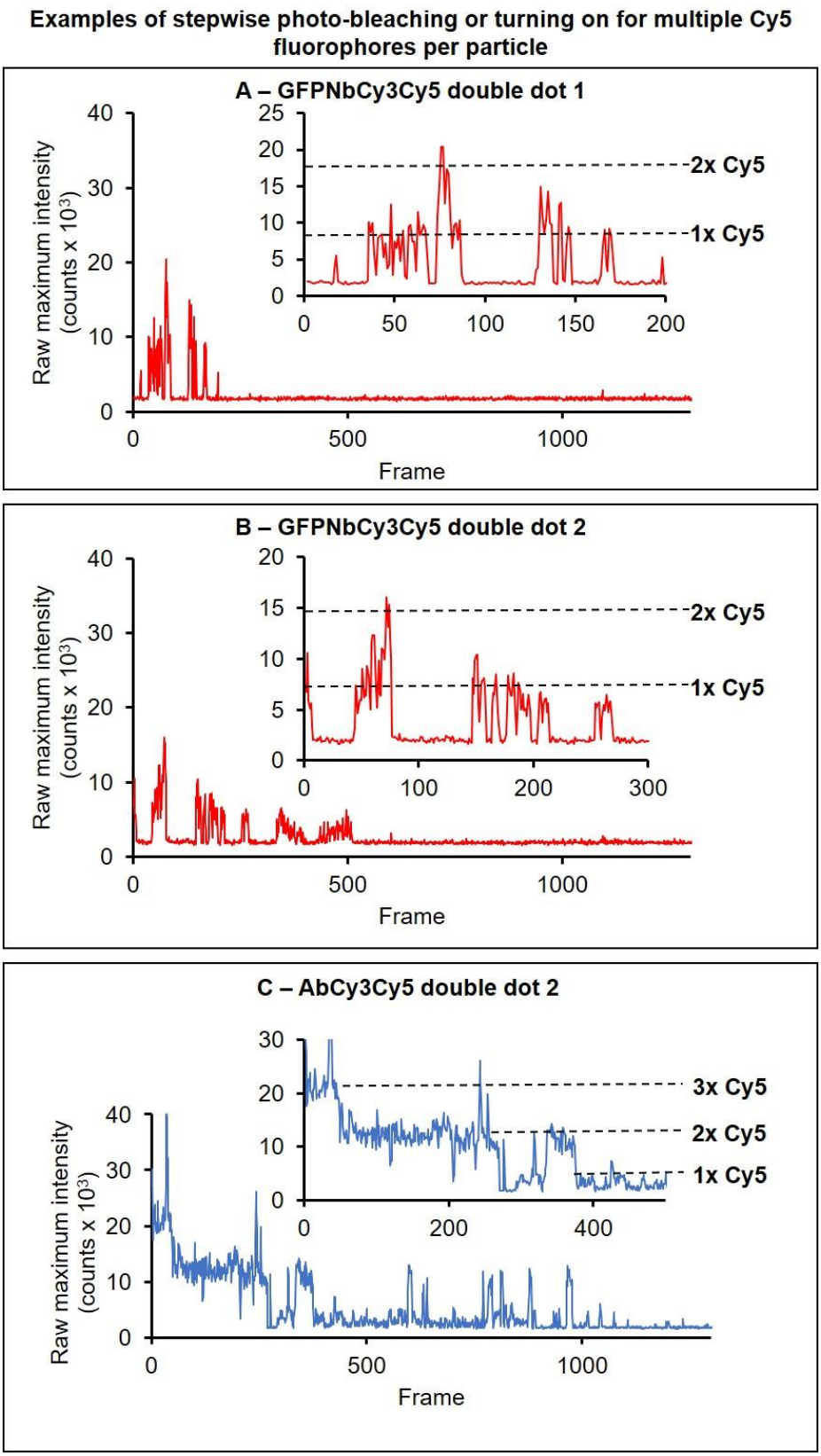


Figure 4.14 – Stepwise photobleaching and activation for single particles. *Maximum intensity per frame over time for single particles of* ***A*** *and* ***B*** *GFPNbCy3Cy5 and* ***C*** *AbCy3Cy5 particle, with inset zoomed areas on the instances of stepwise photobleaching or activation. Suggestions of numbers of Cy5 fluorophores per particle are shown with dashed lines.*

4.2.6 Single AbCy3Cy5 particles show heterogeneity

The same analysis of single particles was carried out for AbCy3Cy5, which was thought to be labelled randomly with a Poisson distribution. Multiple Cy5 fluorophores per particle should have been visible in some cases and **Figure 4.15** shows the results. Detected photons histograms for four particles in **Figure 4.15A** and **B** showed distributions from 200-4500 photons per localisation, much broader than those evident for any GFPNBCy3Cy5 localisations, they also showed multiple peaks.

Particles **B** and **D** mad a modal number of detected photons of 300-500, whereas, and **A** and **C** had a modal value in the range 800-1200. As previously discussed this could have been the result of multiple fluorophores per AbCy3Cy5 particle, and the GFPNbCy3Cy5 data suggested one Cy5 fluorophore could have ~600-700 photons detected per localisation although there were dimmer fluorophores present as well. These data and the presence of lower peaks in **Figure 4.15A** and **C** indicate that 800-1200 photons per localisation is probably the result of two Cy5 fluorophores emitting simultaneously. This was strongly supported by the stepwise photobleaching trace observed in **Figure 4.14C** which clearly show multiple fluorophores are present. This was the same particle as **Figure 4.15C**, which helps explain the large modal number of photons per localisation.

Therefore, the number of photons detected per single Cy5 fluorophore for the AbCy3Cy5 construct was ~300-500, in broad agreement with the GFPNbCy3Cy5 data (600-700). The difference may have been due to the physical differences of the constructs affecting photophysics. It was also evident from the AbCy3Cy5 photobleaching trace and the multiple peaks in the intensity histograms above 1000 photons that multiple Cy5 fluorophores were observed per particle, indeed particles **A**, **C** and **D** all showed good evidence of at least three Cy5 fluorophores, whereas trace **B** had too few localisations at high intensity to be obvious. Particle **C** may have contained up to ten Cy5 fluorophores.

These data clearly indicated multiple labelling of each Ab particle with Cy5. This was expected with random lysine labelling, and the average number of Cy5 fluorophores per Ab was measured as 0.88 on average, which theoretically meant 36.5% of particles were singly labelled, 16.1% with two Cy5, 4.7% with three and so on. It was possible to estimate these using intensities of 300 particles. Based on the a single Cy5 emitting 300-600 photons per frame, approximately 4% of localisations contained one Cy5

(4593 out of 115494), thus 96% have more than one. This was not very close to the predicted 36.5%, which may be due to systematic error in the measurement of the average number of Cy5 fluorophores from the protein and dye absorption spectra or that a Poissonian distribution did not describe the labelling in this instance. Nevertheless, this qualitative value of 4% demonstrated that most Ab particles were multiply labelled with Cy5. Indeed, the calculation may have underestimated the degree of multiple labelling because it only considered localisations which contain two simultaneously emitting Cy5 fluorophores, and it may have been that two Cy5 on the same particle do not emit simultaneously. This degree of multiple labelling, makes quantification with this construct very difficult in comparison to the GFPNBCy3Cy5, which had a much more consistent detected intensity.

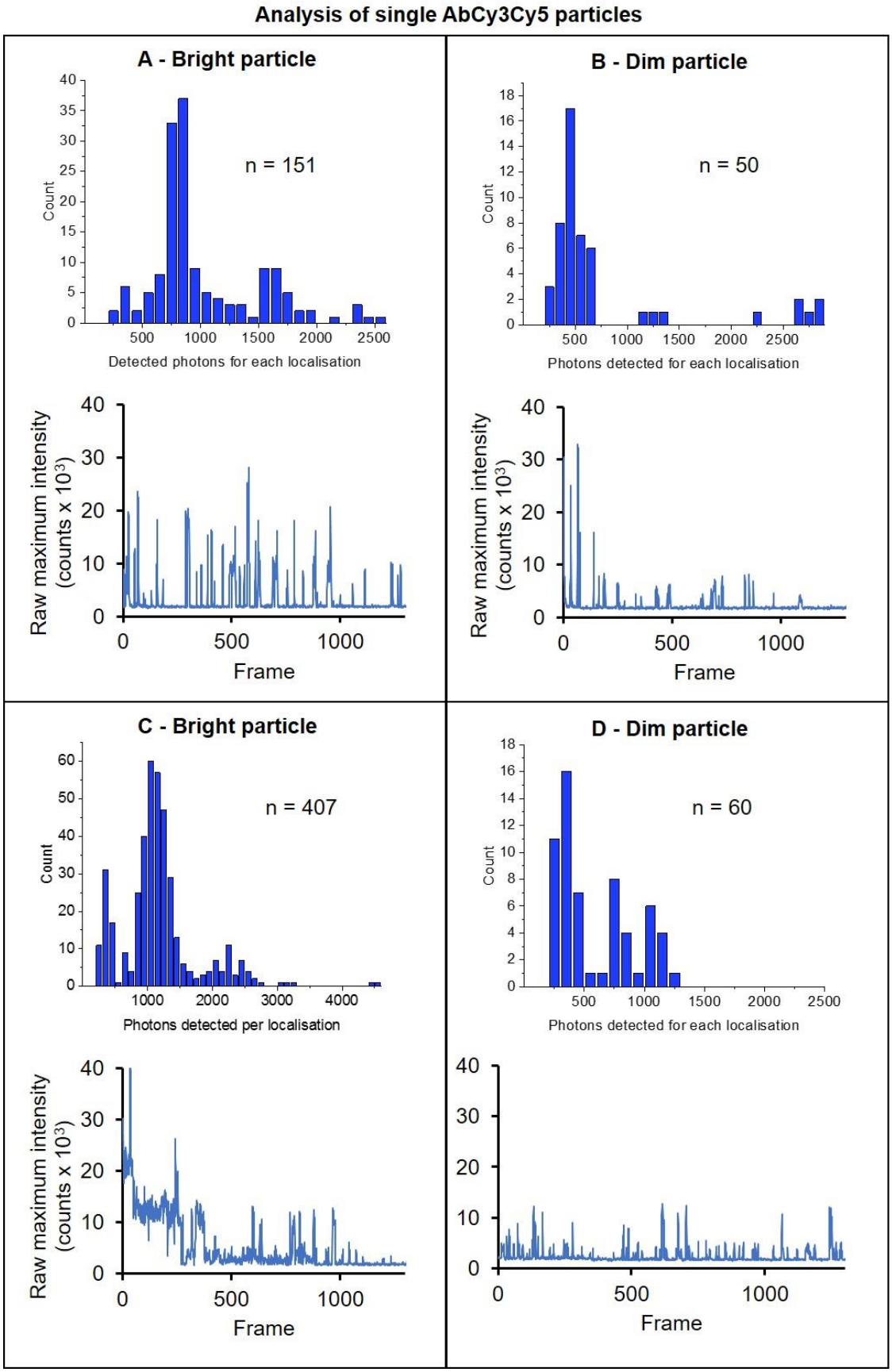


Figure 4.15 – Detected intensity varies between single AbCy3Cy5 particles. *Four single particles* ***A-D have*** *intensity histograms of photons detected and maximum intensity per frame traces. The number of localisations from each trace (n) is shown.*

4.2.7 Photophysical comparison of GFPNbCy3Cy5 and AbCy3Cy5 for quantitative SMLM

The GFPNbCy3Cy5 was designed to be one-to-one labelled with GFPNb-activator and emitter dyes. It was hoped that quantitation should be more accurate and precise with specific labelling ratios compared to the average labelling of Abs. To quantify whether this construct was measurably better for quantitation than a randomly labelled Ab, two constructs with the same activator and emitter dyes (Cy3 and Cy5) were compared directly using the data previously acquired.

Several photophysical parameters were measured for both constructs: the number of photons emitted per localisation, the total on-time, total number of photons emitted per particle, the number of blinks per particle and the number of localisations per particle. All parameters were expected to show a difference between the two constructs, because some AbCy3Cy5 particles appear to have more than one Cy5 fluorophore per particle. The results are shown in **Figure 4.16**. All Cy5 activation events were considered, not only those after an activator pulse.

The number of photons detected for each localisation is one measure of use in quantification. Since AbCy3Cy5 can contain multiple Cy5 fluorophores per particle, in some cases two Cy5 will be on during the same frame and detected as the same particle because they are very close, which should result in a greater number of photons detected for that particle, in that frame. In the GFPNbCy3Cy5 case, there should only be one Cy5 per particle, so the number of detected photons should be a tight Gaussian with the average number of photons for one Cy5. **Figure 4.16A** shows the number of photons detected per localisation for both constructs. The GFPNbCy3Cy5 shows on average a lower number of photons detected and a much tighter distribution than the AbCy3Cy5 construct, indicating that AbCy3Cy5 has a measurably higher number of frames in which two or more Cy5 are simultaneously fluorescing.

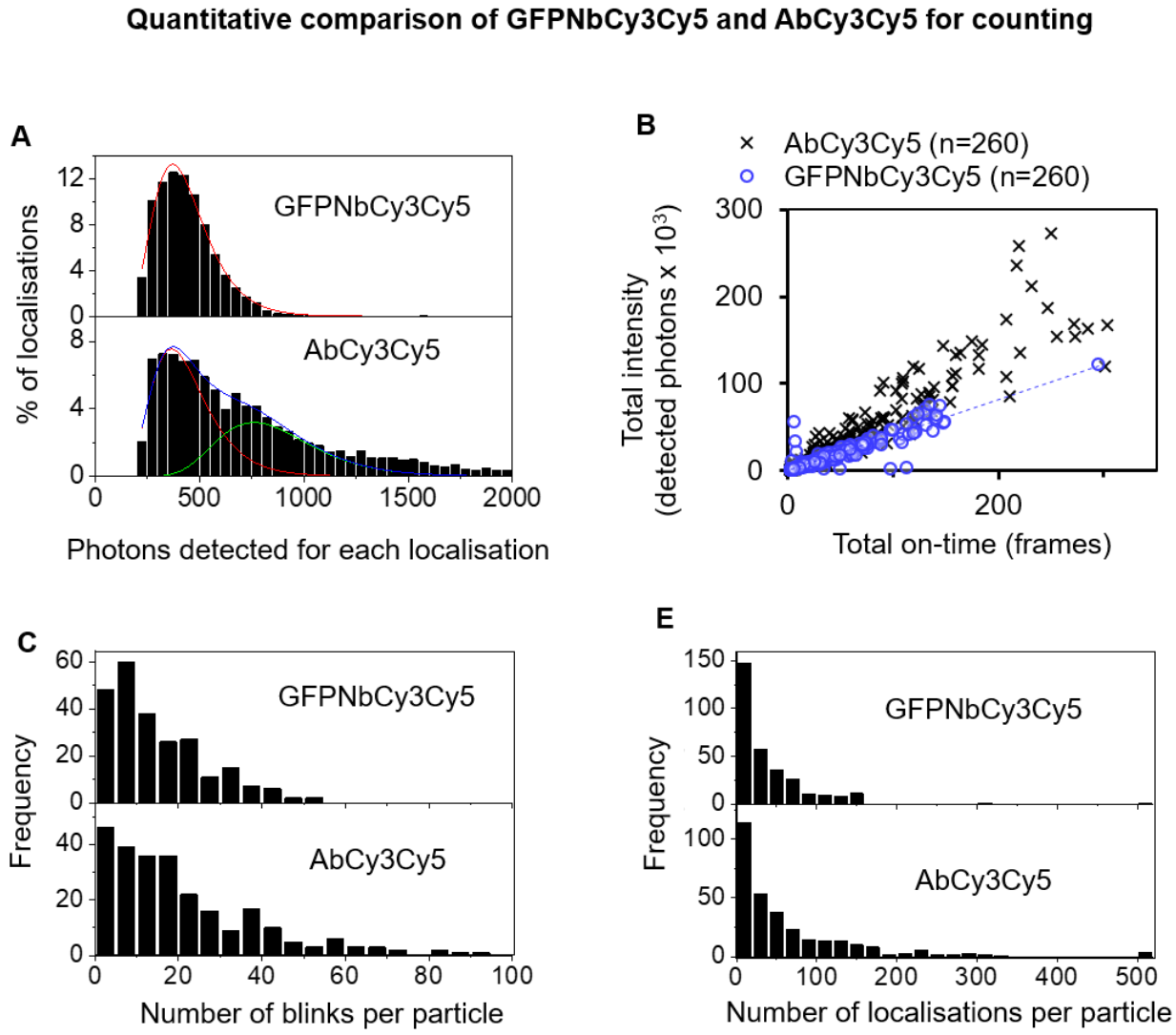


Figure 4.16 – Quantitative comparison of GFPNbCy3Cy5 and AbCy3Cy5. A *Histograms of photons detected per localisation excluding localisations that are the first or last of a blink. The coloured lines represent log-normal curves fit to the data, GFPNbCy3Cy5 was fit to a single, whereas, AbCy3Cy5 was fit to two log-normal functions (red and green), the overall function was plotted in blue.* ***B*** *Total on-time compared to total detected photons per each particle.* ***C*** *Histograms of number of blinks per particle as used in previous publications for counting* (Hummer, Fricke, and Heilemann 2016; Karathanasis et al. 2017; Krüger et al. 2017)*.* ***D*** *Histograms of Number of localisations per particle, similar to the counting method used in the Lakadamyali group* (Ricci et al. 2015; Zanacchi et al. 2017)*. All 4 charts were compiled using the same two datasets filtered for tracks that are longer than one frame, and all histograms were plotted with bin width close to optimal as stated by (Izenman 1991).*

For counting experiments, it is useful to have consistent fluorophores that behave in a predictable manner. One way to measure how consistent GFPNbCy3Cy5 was compared to AbCy3Cy5 was to compare the total number of photons and total on-time for single particles. It was expected that GFPNbCy3Cy5 would have a more consistent total number of photons and total on-time because only one Cy5 could ever be detected and the photon budget for Cy5 is constant under the same conditions. AbCy3Cy5 however, appeared to have multiple Cy5 fluorophores detected in the same

frame. Two Cy5 fluorophores in one particle would not have twice the total on-time because they are both imaged at the same time and their on-times seem as though they overlap. AbCy3Cy5 was, therefore, expected to have a larger number of total detected photons compared to the total on-time. The total number of detected photons was compared to the total on-time for the same particles. **Figure 4.16B** shows the total detected photons plotted against the total on-time for each localisation and for both constructs. GFPNbCy3Cy5 shows very linear relationship that is neatly bounded by ~200 frames and ~100 photons per particle. This indicates a very consistent nature to this construct – it emits a consistent number of photons and bleaches after a consistent number of frames, making it potentially a good tool for counting experiments. AbCy3Cy5 on the other hand shows a less linear relationship with many points above the GFPNbCy3Cy5 line. This suggests considerable heterogeneity compared to GFPNbCy3Cy5. Also, the linear part of the AbCy3Cy5 data has a slightly higher gradient (550 photons per frame) than GFPNbCy3Cy5 (400 photons per frame) as calculated from lines of best-fit, which may be a result of multiple Cy5 molecules per particle producing an average higher number of detected photons per frame. This is further evidence to suggest that the AbCy3Cy5 construct has multiple Cy5 fluorophores detected in the same frame in a significant proportion of frames.

One of the reported problems for counting with *d*STORM and STORM is how much the preferred dyes Alexa647 and Cy5 blink. This is partly because the number of blinks per fluorophore is a skewed distribution. A consistent number of blinks would be easily measurable and could then be used to divide total number of blinks observed to obtain a number of fluorophores, as has been done previously (Lando et al. 2012). However, a skewed distribution limits the effectiveness of this method, although fitting number of blinks per fluorophore to known functions has allowed effective counting (Hummer, Fricke, and Heilemann 2016; Krüger et al. 2017). The narrower the distribution of blinks, the more accurately the calibration can be for counting. Randomly labelled Abs contain distributed numbers of fluorophores per particle, making this calibration even more difficult. GFPNbCy3cy5 was expected to have one Cy5 per particle, so the distribution of blinks per particle should be lower than AbCy3Cy5, and this could be advantageous in counting experiments. The number of blinks per particle for both constructs is displayed in **Figure 4.16C**. Broadly speaking the blinking histograms were quite similar for GFPNbCy3Cy5 and AbCy3Cy5, however, the absolute range, inter-quartile range, mean and median values for GFPNbCy3Cy5 were considerably

lower, because AbCy3Cy5 produced a long tail. This was probably because there were multiple Cy5 fluorophores per AbCy3Cy5 particle, which means a larger number of blinks per particle. Twice as many blinks were not observed in most cases because the number of blinks per fluorophore was random and they seem likely to have overlapped by fluorescing in the same frames a significant proportion of the time. The number of blinks did not take into account simultaneous activation of more than one fluorophore in the same spot, which counted as one blink rather than multiple. This GFPNbCy3Cy5 behaviour is indicative of a more predictable STORM probe that is potentially suited for counting experiments, future work could use this sort of data to attempt counting using methods published by the Heilemann laboratory (Hummer, Fricke, and Heilemann 2016; Karathanasis et al. 2017; Krüger et al. 2017; Malkusch and Heilemann 2016).

Another method previously used for quantification of dyes was to count the number of localisations per particle, and use the median combined with a calibration curve (Ricci et al. 2015; Zanacchi et al. 2017). We therefore compared the same constructs in terms of number of localisations to see if GFPNbCy3Cy5 was significantly different to AbCy3Cy5, the data are shown in **Figure 4.16D**. The results were similar to the number of blinks, because the histograms look fairly similar by eye, but GFPNbCy3Cy5 produced a lower absolute range, interquartile range, median and mean, indicating that it is more consistent, and may be better for counting using that method as well. This benefit could be most useful when trying to resolve numbers of fluorophores that are very close or with relatively small numbers of particles, because using the median number of localisations only works with a very large number of particles or where the difference between two samples is large. The greater consistency of GFPNbCy3Cy5 may be beneficial for counting with this method.

These methods of comparing GFPNbCy3Cy5 and AbCy3Cy5 strongly indicate that that there is only one Cy5 fluorophore per GFPNbCy3Cy5 particle, but multiple – up to 5 or more on each AbCy3Cy5 particle, some of which emit in the same frame. The result is that the number of detected photons, total on-time and the number of blinks is measurably less for GFPNbCy3Cy5. This is strong evidence that the one-to-one labelling method has been successful and is measurably superior compared to the AbCy3Cy5 labelling method and may of benefit with different methods of counting.

4.3 Summary

This work has shown that the GFPNbCy3Cy5 STORM probe synthesised previously is functional for STORM because the Cy5 activation is dependent on the power of an activation laser pulse. That activation was shown to be specific to the intended activation laser with very low crosstalk at most activation laser powers. These results were comparable to control Abs randomly labelled with Cy3Cy5 or Alexa405Cy5, as used in previous reports (Ricci et al. 2015; Bates et al. 2007, 2012; Lakadamyali et al. 2012), which indicates that GFPNbCy3Cy5 could be used for multicolour STORM imaging. When a second STORM pair containing Alexa488 and Cy5 was attached to the same GFPNb there was no evidence of activation by the 488 nm laser, so it was deemed to be non-functional for STORM

Quantitative comparison of GFPNbCy3Cy5 and AbCy3Cy5 using single particles and ensemble statistics show that GFPNbCy3Cy5 has fewer Cy5 fluorophores per particle – most likely the designed one-to-one ratio. This results in a narrower distribution of detected photons, total on-time, number of blinks per particle, and number of localisations per particle. This is strong evidence that GFPNbCy3Cy5 is a superior construct for counting, because calibration would be much more precise with that narrower distribution.

Due to time limitations, it was not possible to label the mCherryNb and to attempt either dual colour imaging or quantitation of biological targets, but the collect data indicates that both may be feasible with this construct design.

4.4 Discussion

The investigation into potential multicolour imaging (**sections 4.2.3 and 4.2.4**) demonstrated that the GFPNbCy3Cy5 STORM pair and two controls AbCy3Cy5 and AbAlexa405Cy5 all have Cy5 activated by their respective activator lasers, and exhibit a crosstalk of 0.01-0.3 with other activator lasers in the range 0.5-2 W/cm^2. The GFPNbCy3Cy5 and AbCy3Cy5 constructs were both more specific than the AbAlexa405Cy5 by about ~10-fold. This difference was probably due to different imaging conditions, which were improved by increasing read-out laser power and optimising illumination, for the later imaging of the Cy3Cy5 constructs, so this may not be a real difference. Published reports have shown crosstalk of <10% for Alexa405Cy5 and <10% for Cy3Cy5 (Bates, Blosser, and Zhuang 2005; Bates et al. 2007), so the all three constructs seem with the acceptable range for multicolour imaging, although in some cases the Alexa405Cy5 crosstalk has been observed to be immeasurably low (Bates et al. 2007). Whilst it is important to have as low a crosstalk as possible, these values should be good enough for dual-colour imaging, as previously reported. To improve the crosstalk for improved channel un-mixing it would be advantageous to try to lower the crosstalk of the AbAlexa405Cy5 construct – this could be done by optimising illumination conditions for this construct too.

The effect of the 488 nm laser on both Alexa405Cy5 and Cy3Cy5 was noted to be considerable, and this also reflects previous reports of crosstalk of 10-20% (Bates et al. 2007, 2012). This result is probably why recent dual colour STORM studies prefer the combination of Alexa405Cy5 and Cy3Cy5 rather than Cy2 or Alexa488 as activators (Ricci et al. 2015).

For fixed cell imaging, however, new controls are necessary because crosstalk will probably be different in HILO illumination required for cell work. In some cases, the crosstalk has been measured from the real sample of interest (Bates et al. 2012), because the STORM pairs were spatially separated, allowing for internal controls. This was effective because it accounts for any change in setup that may accidentally occur between control and real samples, but is unlikely to be possible in co-localisation experiments, for example.

Published STORM work has used predominantly randomly labelled Abs (Ricci et al. 2015; Zanacchi et al. 2017; Bates et al. 2007, 2012). The random nature of the labelling leads to quantification artefacts, which have been qualitatively observed here. They

broaden distributions of detected intensity, blinks and total on-time for single particles, whereas, the GFPNbCy3Cy5 construct here has much narrower distributions. The important role of one-to-one characterised here may allow quantitative imaging, and may also be combined with multicolour imaging, for additional benefits. The results described here are also likely to apply to the previously reported singly labelled Nbs developed by Plenier et al (2015), although they were only used for *d*STORM.

The photophysical characterisation carried here clearly shows that the GFPNbCy3Cy5 construct has less variable distributions of all the characteristics measured. However, time limitations prevented attempts at counting molecules in biologically relevant samples. In one respect 'counting' was done by the identification of multiple Cy5 fluorophores in most imaged AbCy3Cy5 particles via the intensity of photons detected, which can in low numbers of molecules be used as a method for counting. However, with numbers above ~5, the Gaussian distribution of fluorescence intensity makes counting very difficult, and this is also true of counting blinks, which has been done with PA-FPs that blink to a low extent (Lando et al. 2012; Annibale et al. 2011), but not for dyes that blink up to ~100 times (Heilemann et al. 2005). One measure that is apparently consistent for the GFPNbCy3Cy5 construct is the total photons and total on-time, which when plotted together produced a very linear distribution that was specifically bounded at 100,000 photons and 200 frames (500 photons from every frame). It may be that this parameter could be most used to count the number of fluorophores in a diffraction limited area, especially since it is more likely to have a true Gaussian distribution. Future work could investigate this possibility by counting the number of photons per frame for a particle and dividing it by the consistent ~500 photons per frame observed here.

It is important to note that the analysis of suitability for counting experiments here has been done on all localisations regardless of where they appear during the videos. STORM analysis only assigns channels to particles that are have been activated in at least one specific activation pulse, and so will only use a subset of all localisations. Therefore, the results here may not apply to counting with STORM data. The results here do apply directly to *d*STORM analysis, which would use all localisations, in the same way this analysis has. The results are likely to be the same for STORM but a comparison is necessary to verify this, especially in the possible case of quantitative dual colour STORM, which may complicate the analysis.

This investigation has only used one buffer system for STORM imaging, the conventional GLOX buffer containing an oxygen scavenger system and cysteamine (MEA) to maintain consistency and comparability with previous studies (Ricci et al. 2015; Bates, Blosser, and Zhuang 2005; Bates et al. 2007, 2012). However, alternative buffer systems (Olivier et al. 2013; Aitken, Marshall, and Puglisi 2008) and additives such as COT (Olivier et al. 2013) have been shown to improve photophysical characteristics of cyanines for SMLM. There was insufficient time to investigate the effect of these systems, but it may prove beneficial to characterise their effect on the photophysics for counting and multicolour STORM as well.

5 Discussion and Future work

5.1 Aims and results summary

The main aims of this work were to synthesise one or more STORM pairs that were labelled with activator and emitter dyes in a one-to-one ratio, test successfully synthesised constructs for correct targeting in fixed cells, and for multicolour and quantitative imaging. The one-to-one labelling ratio was thought to be important in allowing more quantitative imaging to be done, because most current STORM work has used randomly labelled Abs (Ricci et al. 2015; Zanacchi et al. 2017; Bates et al. 2007, 2012) and random labelling was difficult to correct for in counting experiments (Ricci et al. 2015).

This work used modified peptides to allow one-to-one labelling with the chemospecific reactions click and NHS chemistry. The peptides were modified with either HaloTag ligands or an SBz group to allow direct labelling of HaloTag fusion proteins or protein targeting modules with an N-terminal Cys for NCL. These were adopted as a flexible method to facilitate multicolour STORM because similar peptides could potentially be labelled with different activator-emitter pairs and targeted to different proteins of interest via HaloTag or immune-derived proteins.

Dual labelled HaloPeptides based on the HaloTag(O2) or HaloTag(O4) ligands were successfully synthesised with a variety of activator-emitter pairs. However, when they were tested for binding in *S. pombe* or mES cells, dual labelled constructs could not specifically bind and covalently link to Halo-tagged target proteins. Singularly labelled Halo(O2)Peptide-sulfo-Cy5 and HaloTag ligand-Atto655 were able to label specifically, so adding a second dye appeared to prevent docking and covalent linking. This was thought to be either a steric hindrance problem or a non-specific charge interaction between the negatively charged dyes and positively charged proteins in cells such as histones because HaloTag ligand-Alexa647 also exhibited some non-specific interactions.

An SBz-peptide was also successfully dual labelled with the same chemical reactions to two different activator-emitter pairs, and then successfully linked to an expressed GFPNb with an N-terminal Cys by NCL. This was also attempted with an mCherryNb, and whilst NCL was partially successful, the expressed protein was post-translationally modified which seemed to prevent complete NCL. Two scFvs were also designed for

NCL but they could not be expressed in the periplasm as required for revealing of the N-terminal Cys in these constructs.

GFPNbCy3Cy5 specifically bound Cnp1-GFP in *S. pombe*, proving it was a fully capable of epitope binding and was also used for SMLM in these cells, showing its applicability to fixed cell imaging. This construct, as well as the equivalent GFPNbAlexa488Cy5 and two Abs labelled in the conventional random manner were assessed for compatibility with STORM on SM surfaces. All except GFPNbAlexa488Cy5 showed a fractional increase with their specific activation lasers, and little to no fractional increase with non-specific activation lasers. The same data was processed using published STORM algorithms (Bates et al. 2007, 2012), which showed low crosstalk. Both measures indicate that these three constructs were good candidates for multicolour STORM imaging. The GFPNbAlexa488Cy5 construct showed no fractional increase with its specific activator laser, so it was concluded that this STORM pair was not functional, and it was not characterised further.

GFPNbCy3Cy5 and AbCy3Cy5 were also compared on SM surfaces for photophysical characteristics that may be useful for effective quantitation. It was found that AbCy3Cy5 surfaces were heterogeneous with many single particles containing more than one Cy5 fluorophore each, up to ten were imaged in some cases as was expected for random lysine labelling. GFPNbCy3Cy5 surfaces on the other hand were more homogenous, with particles containing two Cy5 fluorophores very rare and more than two not observed. This difference in single particles was reflected in averages over ~300 particles, where GFPNbCy3Cy5 showed a tighter distribution in detected photons, maximum intensity per trace, blinks and total on-time. These characteristics are suited to better quantitative imaging than randomly labelled Abs.

With respect to the aims at the start of this project, one-to-one dual labelled activator-emitter constructs for STORM have been successfully synthesised, testing has confirmed that they bind specifically to their targets and work in multicolour and quantitative imaging on glass. Time constraints have prevented these from being extended into multicolour or quantitative imaging in cells, but future work based on what has been learnt here is now feasible.

5.2 Future directions: improving these constructs

These constructs were completely new designs and so it might be possible to improve them with small changes to design or practice.

The HaloPeptides were generally unsuccessful because when dual labelled with Cy3 and sulfo-Cy5 or Alexa488 and sulfo-Cy5 they appeared to bind non-specifically throughout the cells they were tested in. One possible explanation of this is that the negatively charged dyes caused charged interactions with positively charged proteins in the cells, such as histones. To test this hypothesis, the HaloPeptides could be labelled with less negatively charged dyes, for example Cy5 (unsulphonated). Activator-emitter pairs including this cationic dye, might be less inclined to bind non-specifically. Whilst Halo(O4)-Cy5 was synthesised, there was insufficient time to dual label and test that construct, especially with its hydrophobic nature and incompatibility with aqueous solutions. Possibly carrying out the NHS reaction first might allow for a more amenable construct. Similarly, less charged version of the activator dyes might be preferable. There is currently no unsulphonated version of Cy3 available, but Alexa488 has less charge and Alexa405 is more charged. Using the less charged dyes might also be beneficial.

An alternative design strategy might be to use SNAP-tag ligand-based peptides instead of HaloTag ligands. This has been achieved with a SNAP-tag STORM pair with Cy3 and sulfo-Cy5 (Dellagiacoma et al. 2010), where the dyes were both linked using sequential NHS reactions and protective chemical groups in a four-step synthesis (Dellagiacoma et al. 2010 – supporting information). The BG moiety for SNAP-tag contains amines so it was surprising that chemospecific reactions were achieved without compromising ligand binding, although yields of some steps were less than 50% which may reflect this. This was a more difficult synthesis than those carried out here, but it was shown to be effective, and so maybe useful for multicolour quantitative STORM imaging.

Another alternative could be to design the peptides with multiple sites for labelling with dyes. Whilst this project was based on a one-to-one ratio of activator to emitter, the photophysical results suggest that the critical ratio might be the targeting domain to the emitter dye, because having one Cy5 in GFPNbCy3Cy5 narrowed the photophysical distributions compared to the randomly labelled AbCy3Cy5. This does not conclude anything about the number of activators, and Abs are often labelled with more than one dye per Ab on average. Adding reactive lysines for more than one activator may improve some characteristics. For example, it was noted that AbCy3Cy5 had slightly lower crosstalk than GFPNbCy3Cy5 (**Figure 4.11**), which may be a result of fewer activator dyes. Quantification of any effect would be informative.

The SBz-peptide was also successfully labelled, which made it potentially useful for reacting with any targeting protein with and N-terminal Cys. Nbs against mCherry and GFP were used for NCL here, but this could now be extended to other Nbs, scFvs or other suitable proteins. The mCherryNb was not satisfactorily expressed and purified and some work could produce another construct for multicolour imaging there. Also, anti-FLAG and anti-CHD4 scFvs generated by MedImmune were designed with N-terminal Cys for NCL and further investigation may find that they are also suitable for multicolour quantitative STORM. Time limitations prevented investigation into this avenue during this thesis. However, the N-terminal covalent modifications observed for both the GFPNb and mCherryNb may be a general effect of expressing proteins with an N-terminal Cys. To make the method used here generally applicable, it would be desirable to determine exactly what those modifications are and how to avoid them in future constructs.

This work used the strategy of NCL via an N-terminal Cys revealed by signal peptidase cleavage in the periplasm, which required a periplasmic expression system. This is not the only way to generate an N-terminal Cys, and other strategies may be worth investigating. One method that may be used is protease cleavage during purification, common proteases used include TEV and thrombin. Use of these enzymes is usually combined with cytoplasmic expression, which has the benefit of much greater protein yield. Two reports have shown folded Nbs can be expressed in the cytoplasm in *E. coli* (Kubala et al. 2010) and HEK 293T cells (Kirchhofer et al. 2010) whilst others also used periplasmic expression (Saerens et al. 2004; Rothbauer et al. 2006), so seems likely that cytoplasmic expression could be combined with proteolytic cleavage to reveal the N-terminal Cys for NCL.

Another alternative strategy could be to employ maleimide linking of dual peptides to targeting proteins like Nbs. This method has been successfully used to image NPC components at high resolution with *d*STORM, and with minimal epitope-fluorophore displacement producing impressive results, although not attempting quantification (Pleiner et al. 2015) and not with STORM. This approach would require a peptide designed with a maleimide group instead of an SBz or ligand group, and a targeting protein with a reduced surface Cys for reaction. It would also be compatible with NHS and click chemistry since it uses sulphydryl groups that are not reactive towards NHS esters nor azides.

5.3 Limitations of this approach

The approach investigated here has shown several limitations, which might be possible to overcome with small adjustments or need entirely different labelling strategies.

One example was the HaloPeptide construct, which although not rigorously tested with suggestions such as uncharged dyes seems quite likely to not work for STORM at all. This is an important finding for all SMLM laboratories, because it could save considerable time and effort in making similar constructs that cannot bind HaloTag specifically in cells. Indeed, it may have already been found in other laboratories, but has gone unpublished.

The N-terminal Cys approach has also proven to be a limiting factor because periplasmic expression is limited by the number of periplasmic transporters, so produces a modest yield of protein, in our experience ~1 mg per L of culture. The modifications of the GFPNb N-terminal Cys then reduce the yield further because the monomeric fraction was usually modified to a much greater extent and we only labelled dimer GFPNb after gel filtration chromatography, which loses approximately half of the NiNTA purified protein. Indeed, this problem led work with the mCherryNb being abandoned completely. Further optimisation of the N-terminal expression is therefore required, and to investigate what causes these modifications and how to avoid them. Other approaches previously mentioned such as maleimide linkages, or cytoplasmic expression might circumvent this problem completely and allow a simpler purification protocol.

Generally, these protein and SBz-peptide constructs appear to work well, but their synthesis has required too much time, and has resulted in less attention and optimisation on the imaging side of this project. Presumably this is often the case with more complex imaging constructs such as Abs linked to docking strands for PAINT (Jungmann et al. 2014), and requires a more concerted effort from multiple specialists such as synthetic chemists, biochemists and imaging specialists. This time-consuming part of SMLM may be why such complex constructs are not reported more commonly and researchers prefer to use well characterised although possibly less optimal constructs such as randomly labelled Abs. Nevertheless, the optimisation work presented here in design, synthesis, expression and purification should allow more streamlined synthesis of dual labelled peptides and protein targeting domains that could be more quickly used for imaging experiments than previously.

5.4 Future directions: quantitative imaging

This construct has yet to be fully tested for quantitative imaging applications. Previous SMLM work including attempts at quantitative imaging have used various well characterised system as a benchmark, to allow good comparison with other methods, these include DNA-origami structures (Hummer, Fricke, and Heilemann 2016; Karathanasis et al. 2017; Zanacchi et al. 2017), the NPC (Zanacchi et al. 2017; Pleiner et al. 2015; Anna Löschberger, Niehörster, and Sauer 2014; Schermelleh et al. 2008; Szymborska et al. 2013), and nucleosome arrays (Ricci et al. 2015). To quantitatively compare this construct to others and validate the potential benefit of the one-to-one ratio of Cy5 to targeting module one of these systems is necessary. Future work could therefore investigate these systems. DNA-origamis are difficult to target with proteins, although not impossible (Zanacchi et al. 2017), whereas, targeting of the NPC is much simpler, especially since GFP fusions for some constituents are already available (Szymborska et al. 2013), this would suit the GFPNbCy3Cy5 construct, and may be worth investigating, as well as testing the similar Nbs with maleimide linked dyes developed by Pleiner et al (2015), which may be similarly suitable for quantitative imaging.

Computational modelling of blinking is now a powerful method for quantitative imaging (Hummer, Fricke, and Heilemann 2016; Karathanasis et al. 2017). Whilst constructs like GFPNbCy3Cy5 may offer better quantitative imaging compared to randomly labelled Abs and similar constructs, it would be even more effective to combine the best constructs with the best modelling. Currently the GFPNbCy3Cy5 construct and the singly labelled Nbs for *d*STORM (Pleiner et al. 2015) seem as though they provide advantages over other dye labelling methods, so combining them with the methods used for dye blink modelling (Karathanasis et al. 2017) might lead to the best quantitative imaging currently possible.

Quantitation using STORM has been reported twice (Zanacchi et al. 2017; Ricci et al. 2015). Ricci et al (2015) achieved this by measuring the number of blinks in calibration samples of nucleosome arrays with known number of nucleosomes and applied this to imaging nucleosomes in cells, which was not counting absolute numbers of fluorophores. Zanacchi et al (2017) quantified absolute numbers with DNA-origami and NPC imaging. Both reports use Insight3 and other custom software for analysis, so to achieve similar results with the constructs described here would require similar code development and application. Currently, it is not clear whether STORM or *d*STORM is

more useful for quantitative imaging, although both are progressing quickly. Certainly, the application of counting to the multicolour strength of STORM is highly desirable. The constructs developed here have the benefit that they could be adapted for either by adding suitable dyes in the peptide synthesis steps.

5.5 Future directions: multicolour SMLM

The construct has also yet to be used for dual imaging, although all indications show that this should be entirely possible. The Cy3Cy5 and Alexa405Cy5 pair both showed low crosstalk and so a double IF experiments with suitable targets should not be very complicated. Previous multicolour STORM reports have begun with imaging spatially separated target proteins, to prevent crosstalk (Bates et al. 2007, 2012), and SMLM imaging studies often use well characterised systems such as cytoskeletal proteins for this (Malkusch et al. 2012; S. Van De Linde et al. 2011; Heilemann et al. 2008; Bates et al. 2007; Lampe et al. 2012; Dellagiacoma et al. 2010). This would therefore be a good starting point for future work with these constructs, and is likely to succeed with these activator-emitter pairs. Previous work has also carried out crosstalk measurements with only the doubly-labelled sample, provided the two colours are spatially separated (Bates et al. 2007, 2012), this could also be done here and the only variables would then be the laser powers, which must be tuned optimally to get good emitter brightness, and ideal activation with each laser pulse. The 405 nm laser has been observed to over-activate fluorophores in biological samples, and so a compromise between obtaining few enough on-events per frame and high enough power to minimise crosstalk may be challenging. This may also be the case for 561 nm laser.

However, whilst dual colour STORM has been used more widely (Bates et al. 2007, 2012; Ricci et al. 2015; Lakadamyali et al. 2012; Huang, Jones, et al. 2008) three or more colours has rarely been used (Bates et al. 2007, 2012). This may be because of the increased crosstalk observed in this regime when using the 488 nm laser (**Figure 4.11**), which probably results in fewer assignments in the 488 nm and 561 nm channels, and so poorer images. To proceed with multicolour imaging of these constructs requires optimisation of the activation laser powers in each sample, especially the 488 nm laser. Imaging of more than three colours with STORM has not been addressed here and would require much more work to investigate the labelling with compatible emitter dyes Cy5.5 and Cy7, their photophysical properties and an optical setup to allow their use. Future work could investigate this possibility, although

the limiting factor is likely to be targeting constructs for as many colours. The GFPNbCy3Cy5 construct is the only one generated here with satisfactory results, although the mCherryNb could also be used for non-quantitative imaging. Further colours would require further targeting approaches such as the anti-FLAG and anti-CHD4 scFvs or SNAP-tag ligand. Triple-colour STORM imaging is therefore not likely in the immediate future, but dual colour is.

What is also of interest is the targets that maybe chosen for application of dual-colour imaging. It is initially necessary to use benchmark cytoskeletal proteins, but there is little biological interest imaging them. It would be of greater interest to use multicolour imaging to look at nuclear protein interactions. Nucleosomes and RNAPII have been imaged with STORM previously (Ricci et al. 2015) and determined that RNAPII activity causes fewer nucleosomes to cluster into clutches in pluripotent cells compared to differentiated cells. It would be similarly interesting to use dual-colour STORM to investigate complex formation in the nucleus, for example the NuRD and polycomb complexes have many constituents that are only sometimes bound to the complex, these interactions would be interesting to study with STORM.

6 Materials and Methods

6.1 Materials

6.1.1 Chemicals and reagents

Reagent	Supplier
Ampicillin	Melford
Acetonitrile (ACN)	Honeywell
Bacto-agar	Formedium
Chloramphenicol	Duchefa Biochimie
Coomassie stain (Instant Blue)	Expedeon
DTT (Dithiothreitol)	Melford
FPLC Gel Filtration columns and machines	GE Healthcare
Formaldehyde 16% (w/v) methanol-free	ThermoFisher Scientific
HaloPeptides & SBz-peptides	Cambridge Peptides
HaloTag ligands	Promega
Kanamycin	Melford
Milk	Marvel
Miniprep kits	Quiagen

Ni-NTA superflow resin	Qiagen
PBS (137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4)	Oxoid Ltd
protease inhibitor cocktail (IV, EDTA-free)	Roche
PVDF membrane	Millipore
SDS-PAGE Gels (NuPage 4-12& Bis-Tris	Novex, Life Technologies
TCEP (Tris(2-carboxyethyl)phosphine)	Generon
Trifluoroacetic Acid (TFA)	ThermoFisher Scientific
Tris (Tris(hydroxymethyl)aminomethane)	Melford
Tryptone	Formedium
Yeast Extract	Formedium

All other chemicals and reagents were purchased from Sigma-Aldrich.

6.1.2 Enzymes

Enzyme	Supplier
Polymerases Q5 and DreamTaq	NEB
Restriction enzymes:	NEB
SacI, XbaI, NdeI, NotI	
T4 Ligase	NEB
Zymolyase	USBiological life sciences
Gibson Assembly kit	NEB

6.1.3 Plasmids and DNA

Construct	Supplier
pSANG10 vector	A gift from John McCaferty
pSANG10 derivatives containing scFvs	Genscript
pOPIN-E containing mCherryNb	Genscript
pOPIN-E vector containing initial GFPNb	Collins laboratory via Addgene
DNA primers	Eurofins

6.1.4 Fluorophores

Fluorophore	Supplier
Alexa488-NHS	Invitrogen
Cy3-NHS	Lumiprobe or GE Healthcare
Sulfo-Cy5- N_3	Lumiprobe
Cy5- N_3	Lumiprobe
Alexa405-NHS	ThermoFisher Scientific

6.1.5 Antibodies

Antibody	Supplier
rabbit derived C-terminal specific anti-His-tag primary (mouse, 121M4789)	Sigma-Aldrich
Anti-mouse IgG (RMG07, monoclonal)	Abcam
Mouse primary IgG (ab80892, polyclonal)	Abcam
Anti-Rabbit IgG (RMG02, monoclonal)	Abcam
Rabbit primary, anti-βtubulin (ab179153)	Abcam

6.1.6 *E. coli* strains

Strain name	Genotype (supplier)
DH10^{TM}	*F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ- rpsL nupG / pMON14272 / pMON7124* (Invitrogen)
BL21(DE3)pLysS	(F- *ompT hsdS*B (rB-mB-) *gal dcm* (DE3) pLysS (CamR); Invitrogen)
LEMO21(DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/ pLemo*(Cam^{R}) *λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5* pLemo = pACYC184-*PrhaBAD-lysY* (NEB)

6.1.7 *S. pombe* strains

Strain name	Genotype (supplier)
Mis6-Halo	*h+ ade6-210 leu1-32 ura4-D18 mis6-Halo*
Cnp1-GFP	*h- leu1 ura4 lys1 GFP-cnp1::NAT* (Takayama et al. 2008)

6.2 scFv expression in *E. coli*

6.2.1 Transformation

The transformation protocol was the same for all constructs that were already characterised. Approximately 1 ng of plasmid DNA was added to 50-100 of either the LEMO21 or BL21 strains whilst thawing on ice and incubated for approximately 30 minutes. They were then heat shocked in a water bath for 45 seconds at 42^0C, before being returned to ice for 2 minutes to recover. 1 ml of LB (Lysongeny Broth, Bertani (1951)) medium was then added, and the mixture incubated at 37^0C for approximately 1 hour. The cells were pelleted by centrifugation at 3,000g for one minute, and 900 µl of the medium removed. The remaining 100 µl was then spread on an LB-agar plate with the appropriate antibiotic added for selection. Plates were incubated overnight at 37^0C to allow the colonies to grow.

6.2.2 LB culture incubation

Starter cultures were grown up overnight by either by picking one colony from transformation plates, or by adding 1 mL of glycerol stock to 10-20 mL LB plus respective antibiotics (kanamycin 50 µg/mL for pSANG10 selection, and chloramphenicol 34 µg/mL for pLEMO selection) and rhamnose at 0.5 mM for the LEMO21 strain, and incubating whilst shanking overnight at 37^0C.

These cultures were grown until stationary phase, and were then diluted 50-fold in either 50 or 500 mL flasks of LB plus antibiotics (plus varying rhamnose concentrations for LEMO experiments). These were also left shaking at 37^0C.

6.2.3 IPTG induction

Cultures were then grown until OD (optical density) 600 = 0.3-0.4, at which point they were induced with 0.4-1 mM IPTG (isopropyl-β-D-1-thiogalactopyranoside), and grown for 3-5 hr at 16^0C, 30 or 37^0C overnight for.

6.2.4 Periplasmic extraction

The periplasmic fraction was extracted from the bacterial cultures according to previously designed protocols specifically for scFvs (Martin et al. 2006). The cultures were pelleted at 3,000 g for 10 min, re-suspended in $1/20^{th}$ the starting volume of 30 mM tris pH 8.0; 1 mM EDTA; 20% sucrose (w/v) plus 25 U/mL benzonase, 0.1 mL/mL protease inhibitor cocktail IV and 1.5 kU/mL lysozyme. After 10 minutes on ice, the cells were pelleted at 3,000 g for 10 minutes, the supernatant decanted and kept on

ice. The remaining pellet was re-suspended in 1/20th the starting volume of 5 mM $MgSO_4$ plus 25 U/mL benzonase, 2 µL/mL protease inhibitor cocktail IV and 1.5 kU/mL lysozyme, and incubated on ice for 10 minutes. The supernatant of this and the previous step (both the periplasm) were pooled. The periplasmic fraction was then spun at 16000 g for 20 min to pellet any remaining cellular debris, and the supernatant decanted and kept on ice.

6.2.5 His-tag purification

The relevant cell fraction was added to ~ 0.5 mL Nickel nitrilotriacetic acid (Ni-NTA) beads, which had been pre-equilibrated in 1x PBS (137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4) with 10 mM imidazole and mixed end-over-end overnight at 4^{0}C. The beads were then washed in three column volumes of 20 mM imidazole in 3x PBS pH 8.4 several times; and the protein eluted in 3-10 mL of 400 mM imidazole, 500 mM NaCl, 50 mM tris, and pH 8.4.

6.3 Nanobody protein expression in minimal medium

Transformations were carried out in the same way as for the scFv expression protocol.

6.3.1 Starter Cultures

Starter cultures were grown in 50 mL LB with 100 µg/mL for pOPIN-E selection and chloramphenicol 34 µg/mL for pLEMO selection overnight or until saturation.

6.3.2 Incubation in minimal medium

The whole 50 mL starter culture was added to 1 L of filter sterilised minimal medium, containing 42.3 mM Na_2HPO_4, 22 mM KH_2PO_4, 9.3 mM NH_4Cl, 9.1 mM NaCl 1 mM $MgSO_4$, 300 µM $CaCl_2$ with trace quantities of EDTA, $FeCl_3$, $ZnCl_2$, $CuCl_2$ $CoCl_2$, H_3BO_3 $MnCl_2$, and had 0.4% glucose, 100 µg/mL ampicillin, 34 µg/mL chloramphenicol plus 1 µg/mL biotin and thiamine added immediately before the cells (1 mL 1000x biotin & thiamine). 1 L cultures were grown whilst shaking at 37°C for approximately 7 hours or until OD600 ~ 0.7-0.8, when the cells were induced with 2 mM IPTG and the incubation temperature lowered to 16°C and grown for a further 2-4 hours.

6.3.3 Media extraction

Cells were removed by centrifugation at 3,000g for 15 min using a JLA10.5 rotor and Beckman-Coulter bucket centrifuge, and the supernatant was decanted and saved. 1 mL samples of the cell pellet and supernatant were taken for subsequent SDS-PAGE analysis. The cell pellets were then discarded. The supernatant was then spun again

at higher 15,000g for 30 min using the same rotor to remove cellular debris and large complexes.

6.3.4 His-tag purification

The supernatant was then applied to ~5-10 mL Ni-NTA superflow beads in a gravity column, which had been pre-equilibrated in 1x PBS (137 mM NaCl, 3 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4) with 10 mM imidazole. The supernatant was allowed to flow through the column, with no additional incubation time. After the whole supernatant had passed over the beads, they were washed in ~50 mL of 20 mM imidazole in 3x PBS pH 8.4. The protein was then eluted in 3-10 mL of 400 mM imidazole, 500 mM NaCl, 50 mM tris, and pH 8.4. The samples eluent fractions were pooled, concentrated into ~ 0.5 mL with a 5 KDa molecular weight cut off (MWCO) spin columns and buffer exchanged into 0.2 M K_2PO_4/$KHPO_4$, pH 7.0.

6.3.5 FPLC Gel Filtration

The protein was then purified by gel filtration on an S75 24 mL Sephadex column using an Akta Explorer (GE Healthcare), with 0.2 M K_2PO_4/$KHPO_4$ pH 7.0 running buffer. The fractions for each peak were then pooled as appropriate and concentrated into ~0.5 mL using 5 KDa MWCO spin columns. As soon as possible the purified protein was used for NCL and stored in 0.2 M K_2PO_4/$KHPO_4$ pH 7.0 containing 2.5 mM TCEP at 4°C.

6.4 Protein Analysis

6.4.1 Western Blotting

A standard SDS-PAGE gel (4-12% bis-tris; Novex, Life Technologies) was run at 200 V for ~40 min and transferred to a Polyvinylidene fluoride (PVDF) membrane in transfer buffer (192 mM glycine, 25 mM Tris, 20% methanol (v/v)) at 150 mA for 90 min. This was blocked using 5% milk in PBS at room temperature (23°C) for 1 hour before primary staining using a rabbit derived C-terminal specific anti-His-tag Ab (Sigma-Aldrich, mouse, 121M4789), shaking for 1 hour in PBS, and secondary Ab staining with an anti-rabbit secondary Ab from goat, conjugated to peroxidase. It was then washed 3 times, 10 min each, shaking, in PBS. Bands were visualised by incubating the PVDF membrane in 0.1 M Tris-HCl pH 8.6 with 0.0006% hydrogen peroxide v/v, 67 micromolar (μM) p-coumanric acid, and 1.26 mM luminol for 1 min and exposing the membrane plus film for 1-600 s and the film developed. For the images in **Figure 3.7** the developed film was placed on top of the membrane to visualise the pre-stained

protein markers relative to the His-tag stained bands. The positive control for the Blot was p48 (48 KDa) with a C-terminal His-tag.

6.4.2 MALDI-MS

MALDI-MS was carried out on site at the protein and nucleic acid chemistry facility at the Department of Biochemistry. The work was done by Dr Len Packman and Dr Mike Deery and his team. For proteins analysis was done on ~10 μM samples in aqeous solution, for labelling peptides analysis was carried out on similar concentrations dissolved in DMSO.

6.5 DNA cloning

6.5.1 Cloning of scFvs

Both scFv constructs were purchased from Genscript as plasmids containing the scFv ORF with the PelB signal sequence. To maximise the chances of good expression the PelB signal sequence was also substituted for the DsbA signal sequence to maximise the chance of it expressing correctly. To do this the PCR primers shown in **Table 6.1** and **6.2** were used to synthesise a small fragment that was cloned in using conventional NdeI and NotI digestion of PCR product and plasmid followed by T4 ligation. Correct clones were selected by DNA Sanger sequencing.

Forward primer (to insert DsbA signal sequence, with NdeI site)	ATACCATATGAAAAAAATCTGGCTGGC GCTCGCGGGTTTGGTTTTAGCCTTCA GCGCATCAGCTTGCGAAGTGCAGCTG CAGCAGAGC
Reverse primer (containing NotI site)	ATGATGTGCGGCCGCGCGTTT

Table 6.1 – Primers used to clone the DsbA signal sequence into the anti-FLAG scFv to replace the PelB sequence in the purchased plasmid. *Orange indicates the DsbA signal sequence, green indicates the N-terminal Cys and Glu residues, bold indicates the start codon and underlined are the NdeI and NotI sites used for cloning.*

Forward primer (to insert DsbA signal sequence, with NdeI site)	ATACCATATGAAAAAAATCTGGCTGGC GCTCGCGGGTTTGGTTTTAGCCTTCA GCGCATCAGCTTGCGAAGTGCATCTG GTGGAAAGC
Reverse primer (containing NotI site)	ATGATGTGCGGCCGCGCGTTT

Table 6.2 – Primers used to clone the DsbA signal sequence into the anti-c-Myc scFv to replace the PelB sequence in the purchased plasmid. *Orange indicates the DsbA signal sequence, green indicates the N-terminal Cys and Glu residues, bold indicates the start codon and underlined are the NdeI and NotI sites used for cloning.*

6.5.2 GFPNb cloning by Gibson Assembly

The GFPNb gene was obtained from the Collins laboratory via Addgene as expressed in Kubala et al (2010). To introduce PelB or DsbA signal sequences at the N-terminus of the GFPNb ORF, as well as the signal peptidase site that would reveal an N-terminal Cys, the Gibson Assembly cloning method was used. For each signal sequence, two PCRs were carried out using the GFPNb-pOPIN-E vector as the template; one PCR product contained the sequence to be added to the ORF and one to introduce a XbaI cloning site upstream of the ORF. Both PCRs were designed to produce the required homology region for Gibson Assembly, the primers for both are shown in **Table 6.3** and **Table 6.4**.

The two PCR products were combined by Gibson Assembly into one long DNA fragment, then cut with restriction enzymes XbaI and SacI. The parent plasmid was also cut with these enzymes, and conventional T4 ligation was used to insert the long fragment. The ligation was transformed into DH10 cloning cells and individual colonies were screened for successful addition of the signal sequence by Sanger DNA sequencing.

Forward primer 1 (to insert DsbA signal sequence, with the a XbaI cleavage site)	ATGAAAAAAATCTGGCTGGCGCTCGCGGGTTTGG TTTTAGCCTTCAGCGCATCAGCTTGCGAAGTTCAA CTGGTGGAAAGCGG
Reverse primer 1 (containing SacI site)	GATGGTGATGGTGATGTTTAGAGCTCACCGT
Forward primer 2 (containing XbaI site upstream of ORF)	TCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGC TA
Reverse primer 2 (to insert homology region)	GCCAGCCATTTTTTTCATGGTATATCTCCTTTGATT GT

Table 6.3 – Primers used for Gibson Assembly of GFPNb to add the DsbA signal sequence and N-terminal Cys. *Orange indicates the DsbA signal sequence and in which the homology region for Gibson Assembly falls, green indicates the Cys and Glu residues. Underlined nucleotides indicate the XbaI and SacI restriction sites, and bold indicates the start codon.*

Forward primer 1 (to insert PelB signal sequence, with the a XbaI cleavage site)	ATGAAATACCTGCTGCCGACCGCTGCTGCTGGTCTGCTGCTCCTCGCTGCCCAGCCGGCCATGGCCTGCGAAGTTCAACTGGTGGAAAGCGGC
Reverse primer 1 (containing SacI site)	GATGGTGATGGTGATGTTTAGAGCTCACCGT
Forward primer 2 (containing XbaI site upstream of ORF)	TCTGGCGTGTGACCGGCGGCTCTAGAGCCTCTGCTA
Reverse primer 2 (to insert homology region)	GTCGGCAGCAGGTATTTCATGGTATATCTCCTTTGATTGT

Table 6.4 – Primers used for Gibson Assembly of GFPNb to add the PelB signal sequence and N-terminal Cys. *Orange indicates the PelB signal sequence and in which the homology region for Gibson Assembly falls, green indicates the Cys and Glu residues. Underlined nucleotides indicate the XbaI and SacI restriction sites, and bold indicates the start codon*

6.5.3 DNA sequencing

All DNA Sanger-sequencing was carried out by the in-house DNA sequencing facility at the Department of Biochemistry. T7 sequencing primers were used in all cases.

6.6 Dye-labelling chemistry

6.6.1 Peptide labelling – click chemistry

Equimolar proportions of Cy5-N_3 and peptide substrate (containing the alkyne group) were mixed together with water, 50% DMSO, 0.2 M triethylammonium acetate (TEAA), and freshly made 0.5 mM ascorbic acid as per the manufacturer's instructions. This mixture was de-oxygenated by bubbling pure nitrogen gas through for two minutes, then 0.5 mM TBTA-Cu(I) complex was added and the mixture again de-oxygenated. The reaction was left overnight at room temperature, and the resulting mixture separated by RP-HPLC.

6.6.2 Peptide labelling – NHS chemistry

Reactions of 25-200 µL were made up in 100 mM $NaCO_3$/Na_2CO_3 and the non-NHS ester reactant was added next at µM-mM concentration. The pH was then altered using 1 M HCl, and measured using a small volume pH probe (VWR). The desired pH varied depending on the experiment. The NHS-ester was then added and the pH could again be changed if desired. The reaction mixture was then incubated at 24°C for 1-12 hours before separation of products by RP-HPLC. The pH was often measured after each individual reagent was added to observe the effect each has on pH, as well at the end of the reaction. Various concentrations of DMSO and pH were tested throughout this work.

6.6.3 RP-HPLC

Purification of the conjugated fluorophore-peptides was carried out using an Agilent 1100 series HPLC machine equipped with a two-position valve (Rheodyne), 100 ul loop (Sigma-Aldrich), a C-18 Security Guard column and a C-18 separation columns (both Brownlee).

Reaction mixtures of either 50 or 100 µL were loaded onto the loop and injected onto a Security Guard C-18 reverse phase column via the two-position switching valve. Elution was achieved with a gradient of 0.1% trifluoroacetic acid (TFA) in acetonitrile (ACN) from 0 to 80% over 20-30 minutes, against 0.1% TFA in water. The precise nature of the gradient was tailored to the reaction mixture being purified.

6.6.4 Native Chemical Ligation

His-tag purified anti-GFPNb was reacted with RP-HPLC purified SBz-peptide with one or two dyes covalently attached using 0.2% (v/v) benzenethiol as a catalyst, at various molar ratios, 1:1 or with labelled peptide in slight excess proved most effective for

maximum labelling. 25 µL reactions were left at room temperature stirring using small magnetic stirrers for ~18-24 h. After the desired period the reaction was halted using 2 mM TCEP (final concentration). To separate the labelled Nb away from unreacted peptide the mixture was made up to 200 µL with Dulbecco's PBS (DPBS) and 2.5 mM TCEP. This was added to a NAP-5 gravity column, which has been previously equilibrated in DPBS with 2.5 mM TCEP and allow to run until the bed was dry. 500 µL DPBS + 2.5 mM TCEP was then added to allow the mixture to flow into the matrix. And 300 µL more was added to elute the protein into 300 µL fractions from the column but not the unreacted SBz-peptide-dyes. In some cases, the fractions containing labelled protein was exposed to another round of NAP-5 purification. The resultant labelling could be analysed by Coomassie stained SDS-PAGE, and fluorescence using a Typhoon-1000 FLA fluorescence imaging machine (GE Healthcare) as well as MALDI-MS (see **section 6.4.2**). The Typhoon fluorescence data was analysed for efficiency using ImageJ.

6.6.5 Random labelling of Abs with NHS chemistry

Abs were mixed with NHS derivatives of both activator (Alexa405 or Cy3) and emitter dyes (Cy5) simultaneously using a protocol from the Lakadamyali group. To achieve a close to on-to-one labelling with the emitter, a slight excess of ~1.2-fold of the dye was added. For the activator, a larger number of dyes per Ab was necessary to ensure that one was close enough to the emitter to allow STORM, so a ~3-5-fold excess was used. The reaction mixture was left at 24°C for approximately half an hour, and then unreacted and NHS hydrolysed dyes were separated from the protein using a NAP-5 column, with elution in PBS. The absorption at 280 nm, 650 nm and a wavelength appropriate to the activator were used to measure the concentrations of protein, Cy5 and activator respectively, to provide the degree of labelling achieved. Roughly 0.7-1 was achieved with the emitter and 1-3 with the activator, although these were average measurements.

6.7 SM surface imaging, sample preparation and analysis

6.7.1 Slide and sample preparation

Glass slides with 800 µl chambers (Ibidi) were cleaned with either 1 M KOH or 2% Hellmanex and 15 min sonication, before rinsing in MilliQ water. The surfaces were coated with 100 µl poly-L-Lysine and left to settle for 30-60 min, excess liquid was removed, and surfaces dried with compressed air. Respective epitopes such as PA-GFP, or primary Ab were added at variable concentration (10^{-6}-10^{-9}) with 5% BSA which all covered the surface and left to settle for 15-30 min. Finally, ~1 µM labelled Ab or GFPNbCy3Cy5 was applied to the surface with 0.5% BSA and left to settle for 15-30 min before excess was washed away with 20 µm filtered PBS. Imaging buffer containing 10% w/v glucose, 50mg/ml glucose oxidase, catalase and 100 mM MEA in 20 µm filtered PBS as previously described (Sebastian Van De Linde et al. 2011) was added and the chamber sealed without bubbles with a small coverslip. One sample was then imaged for 1-2 hours before the buffer was changed, since the glucose was depleted.

Control slide surfaces were prepared in the same way except that the Ab or GFPNb labelled with fluorophores was not added. The imaging buffer and acquisition of data was the same for each type of surface and experimental samples.

6.7.2 Microscope setup

SR imaging of SM surfaces was carried out on a custom-built Nikon Ti Eclipse microscope that has been previously described (Virant et al. 2017). The sample was illuminated with an OBIS LX 637 nm laser (Coherent) which was filtered through a ZET 640/10 bandpass, modulated by an Acousto-Optic Tunable Filter (Gooch & Housego, TF525-250-6-3-GH18) and focused by a ZET405/488/561/640 m dichroic mirror (Chroma) onto the back focal plane of the objective, resulting in a final intensity of ~300 W cm^{-2} on the sample. Activation of dye pairs was achieved with sequential pulses of 405 nm, 488 nm and 561 nm light (405 nm OBIS, 561 nm OBIS, 730 nm OBIS, 488 nm Sapphire; all Coherent Inc., Santa Clara, California USA). The pulses were controlled by the Acousto-Optic Tunable Filter, triggered every 30 imaging frames. The readout was collected by blocking the laser light by the bandpass ZET405/488/561/640 and passing through a 689/23 nm single-band bandpass filter (All filters AHF Analysentechnik AG), then collected by a CFI Apochromat TIRF ×100 objective with a numerical aperture of 1.49 (Nikon), recorded

on an iXON ULTRA 888 electron-multiplying charge-coupled device (EMCCD) camera (Andor). The pixel size was 129 nm/pixel, the gain due to the EM-CCD was 109.48 ADU/photon, and a frame lasted 50 ms.15 pulses of each of the three activation lasers were recorded for each repetition. The camera, microscope and AOTF were controlled by µManager software (Edelstein, Tsuchida et al. 2014) on a PC workstation.

6.7.3 Data acquisition

SM surface videos were collected with 20000-3000 frames at an exposure length of 40 ms and the activation lasers pulsing once every 20-30 frames depending on the experiments. The 637 nm laser was used at 141.47 W/cm^2 and the activation lasers were varied for power dependence experiments between 0.005 and 70 W/cm^2.

6.7.4 Fractional increase and photophysical analysis with RapidSTORM

Localisation lists were generated by RapidSTORM with a fixed global amplitude threshold of 25,000, which identifies maxima, fits them to a Gaussian and filters those Gaussians using the global amplitude threshold of 25,000 counts per 5x5 pixel square (Wolter et al. 2010, 2012).

The fractional increase upon activator laser illumination was calculated by loading raw localisation lists containing X and Y co-ordinates as well as frame number and fitted intensity, into Origin software (OriginLab, Northampton, MA). Taking the modulus of the frame number for each localisation with respect to the number of frames in one pulse cycle and plotting this number as a histogram produced the number of localisation for each frame in one cycle. The fractional increase was calculated according to **Equation 4.7**.

For photophysical characterisation, RapidSTORM localisation lists were used to track individual particles through time using custom software (Endesfelder group, unpublished), tracked particles were analysed with further custom software; Mistral (Endesfelder group, unpublished). Tracks were combined to form traces for each fluorophore blink and for each particle, which contains multiple tracks. Tracks belonging to particles were manually identified and parameters exported for either single particles or many particles, with the first and last frames of each track removed because they are often dimmer since the fluorophore is only on for part of a frame. Fitted localisation intensity, track length, number of tracks per particle were exported for investigation and graphically analysed using Origin (OriginLab, Northampton, MA).

Histograms were plotted using bin widths approximately equal to the ideal bin width from Feedman-Diaconis – $2(IQR)N^{-1/3}$ (summarised in Izenman 1991), but were the same between compared datasets.

6.7.5 Crosstalk calculations using PeakFit.

Calculation of the crosstalk between different activator channels was done using custom software written by Dr Alex Herbert for PeakFit based on Bates et al (2007, 2012). Raw data was analysed using PeakFit (Herbert 2014). Briefly, the algorithm had noise reduced, local maxima identified, Gaussians fitted and then successfully fitted Gaussians were filtered using SNR = 30, precision = 30 nm, minimum photons = 0 and width factor = 0.5. These resulting localisation lists were stored in memory for use in the pulse activation analysis plugin. Multicolour STORM images were then reconstructed using the crosstalk values calculated and the localisation lists.

6.7.6 Single particle traces and histograms

To compare single particles a single dot was selected from a field of view and a ROI of at least 6x6 pixels was selected around it in ImageJ. The video of that spot was duplicated and checked to ensure there were not more than the one dot appearing throughout the video. For fitted intensity histograms, localisations were generated using a global amplitude threshold of 25,000. The dot was traced using custom software (Endesfelder group, unpublished) and the intensity for each localisation exported using further custom software (Endesfelder group, unpublished). Localisations at the start and end of a blink were removed by a custom written Python script, because the fluorophore may turn on/off halfway through the frame. The resulting localisations had their intensity values plotted into histograms using Origin (OriginLab, Northampton, MA) or into fitted intensity traces with Microsoft Excel.

Raw intensity traces were generated from ImageJ, by filtering the duplicated stack it for the maximum intensity (process, filter, maximum). The intensity of each frame was then plotted against frame number (plot Z-axis profile) and the points exported into Microsoft Excel to plot the graph as desired.

6.8 Labelling and imaging of *S. pombe* and mES cells

6.8.1 Preparation and labelling of *S. pombe* cells for imaging

S. pombe cells were grown in Yeast Extra + Supplements (YES) media from plates or -80°C glycerol stocks at 30°C shaking until OD600 ~ 0.5-1.0. They were then grown for another 24 hours with dilutions in YES keeping them at OD600 <0.6. After 24 hours, the cells were grown to OD600 ~0.5 and 2 mL was spun down into a pellet at 3000g for 2 min. All washes were done in a 1.7 mL Eppendorf tube each with a spin at 3000g for 2 min to re-pellet the cells. The first wash was in 0.5 mL HEMS buffer (20 mM HEPES pH 7.5, 1 mM $MgCl_2$, 1 mM EDTA, 1 M sorbitol), then in 0.5 mL HEMS containing 2.8 mM BME, then the pellet was re-suspended in 0.5 mL HEMS containing 0.25 mg Zymolyase 20T, which was left to incubate at 30°C for 30 min. Subsequently, the suspension was spun at 2000g for 5 min to prevent cell damage in between washes. Pellets were pre-fixed with 0.5 mL HEMS +4% (w/v) formaldehyde for 10 min at 24°C min. Two washes in 0.5 mL HEMS removed the formaldehyde and the cells were then permeabilised with a quick wash of 0.5% Titon-X-100 and EDTA-free protease inhibitor cocktail in HEMS, followed by labelling for at least 1 hour at 24°C in HEMS containing 0.1% tween20, 150 mM NaCl, 1 mg/mL BSA, protease inhibitors, and ~5 pM of labelled anti-GFPNb. After labelling, the cells were washes three times with HEMS containing 0.1% tween20, 150 mM NaCl, 1 mg/mL BSA, EDTA-free protease inhibitors, each for 10 min shaking at 24°C. Post-fixation was carried out in 4% (w/v) formaldehyde for 5 min at 24°C and the fixative was washed once in HEMS, once in PBS and twice in MilliQ (MQ) water. The suspension was then mixed thoroughly by pipetting with a 2 µL tip, which separates clumps of cells. 2 µL was then mounted upon a microscopy dish (Mattek Corporation) which had been treated with poly-L-Lysine and allowed to dry for 30-60 min with excess liquid removed by washing in DPBS. The cells were then left to settle and stick to the poly-L-lysine at 4°C for at least 1 hour. The sample could then be stored for at least 1 week in water at 4°C before imaging. Post-fixed cells not mounted could also be kept for subsequent mounting at 4°C in pure water.

6.8.2 mESC culture and labelling

mES cells transiently expressing CENP-A-Eos3.2-HaloTag were generated by transfecting an adapted version of an CENP-A-mEos3.2 plasmid (Palayret et al. 2015) with HaloTag as well as Eos3.2 (Basu et al. 2018). Cells were grown with maintained pluripotency in suspension with LIF media. For labelling, 35 mm glass bottomed

imaging dishes (Mattek Corporation) were coated in 0.1% gelatin to allow cellular attachment and cells added. After 1-2 days cells were washed once in DPBS and fixed in 4% (w/v) formaldehyde in DPBS for 5 min at 24°C before proceeding to the desired labelling protocol.

Fixed cells were blocked with 1 mL 10% BSA in DPBS for 1 hour at 24°C, before immunofluorescent probing with ~0.1pM-1 nM labelled GFPNb and 10% BSA in DPBS for at least 1 hour at 24°C or overnight at 4°C. Excess probe was washed out with three rounds of DPBS, 0.05% Triton-X-1000 and 0.2% BSA, followed by a wash in DPBS and post-fixation in 4% (w/v) formaldehyde in DPBS for 10 min at 24°C. Excess cross-linker was removed with three washes in DPBS and two further washes in MQ water. These samples could then be kept in MQ water at 4°C for several weeks, if the sample was not allowed to dry out.

6.8.3 Microscope setup and imaging.

All *S. pombe* GFP-Cnp1 and Mis6-Halo and mES cell CENP-A-Eos3.2-HaloTag ensemble imaging was done using a custom-built microscope designed and built by the Klenerman & Lee groups at the Department of Chemistry, University of Cambridge. It was an inverted microscope (IX71, Olympus, Japan) equipped with a 60x oil immersion objective (PlanApo 60x, NA 1.49, Olympus) that resulted in a depth of field of approximately 800 nm and had 4 lasers - a 405 nm (Oxxius, Laserboxx 405, 100 mW); 488 nm (Toptica, iBeam Smart 488 100 mW); 561 nm laser (Cobolt, Jive 200) and 641 nm (Coherent, CUBE 640-100C, 100 mW). Excitation light and fluorescence light were separated using a dichroic mirror (FF410/504/588/669-Di01, AHF, Germany), and appropriate filters were placed in the detection path (488BP, FF01-S20/35; 561 BP; and 641 LP BLP01-635R, all Semrock). The fluorescence signal was recorded with an EMCCD camera (Photometrics Evolve 512). The pixel size was 109 nm with a gain of 43.5 ADU/photon. The instrument was automated with open-source software micro-manager (A. D. Edelstein et al. 2014; A. Edelstein et al. 2010) on a PC workstation.

6.8.4 SMLM imaging in *S.pombe*

GFP-Cnp1 cells were prepared in the same way as for ensemble imaging (**section 6.8.1**) with 1-3 mL with a GLOX scavenger system (Sebastian Van De Linde et al. 2011) added to lysine-coated microscopy dishes (Mattek Corporation), the lid placed

on and sealed with Parafilm. Individual cells were located using white light wide illumination and selected using a small ROI. Before SMLM, images in the white light, GFP and Cy5 channels were taken, and the Cy5 channel with low 641 nm power, was used to focus on one specific centromere cluster for later analysis. Each cell was then videoed with a 641 nm laser individually in small ROI to allow minimum exposure time ~ 10 ms. Initially, the Cy5 fluorophores were turned into a dark state and then began to blink for up to ~70,000 frames. Since the centromere was a very small area ~ 200 x 200 nm, only one fluorophore could blink per frame to allow later localisation. Manual 561 nm pulses at minimum power were therefore used to increase the speed of acquisition. Pulses were kept on for as little time as manually possible (<1 s, probably ~ tens of frames). As activation had less effect later in the video, 561 nm power was ramped via an optical density filter. The powers of the 561 and 641 nm were not measured, as this was only preliminary work, and were likely to vary day to day with a relatively unstable microscope setup.

6.8.5 SMLM analysis

Videos had maxima localised and fitted with RapidSTORM (Wolter et al. 2012)at a global threshold of 25,000, and the localisations were manually selected for the cluster that appeared to be GFP-Cnp1 foci using Microsoft Excel. The selected localisations' intensity was summed for each centromeric cluster, converted from counts to photons and this value was divided by the mean number of photons detected per GFPNbCy3Cy5 from SM surfaces (378.9 photons per GFPNbCy3Cy5). Box plots were then constructed using Origin (OriginLab, Northampton, MA).

7 Appendix

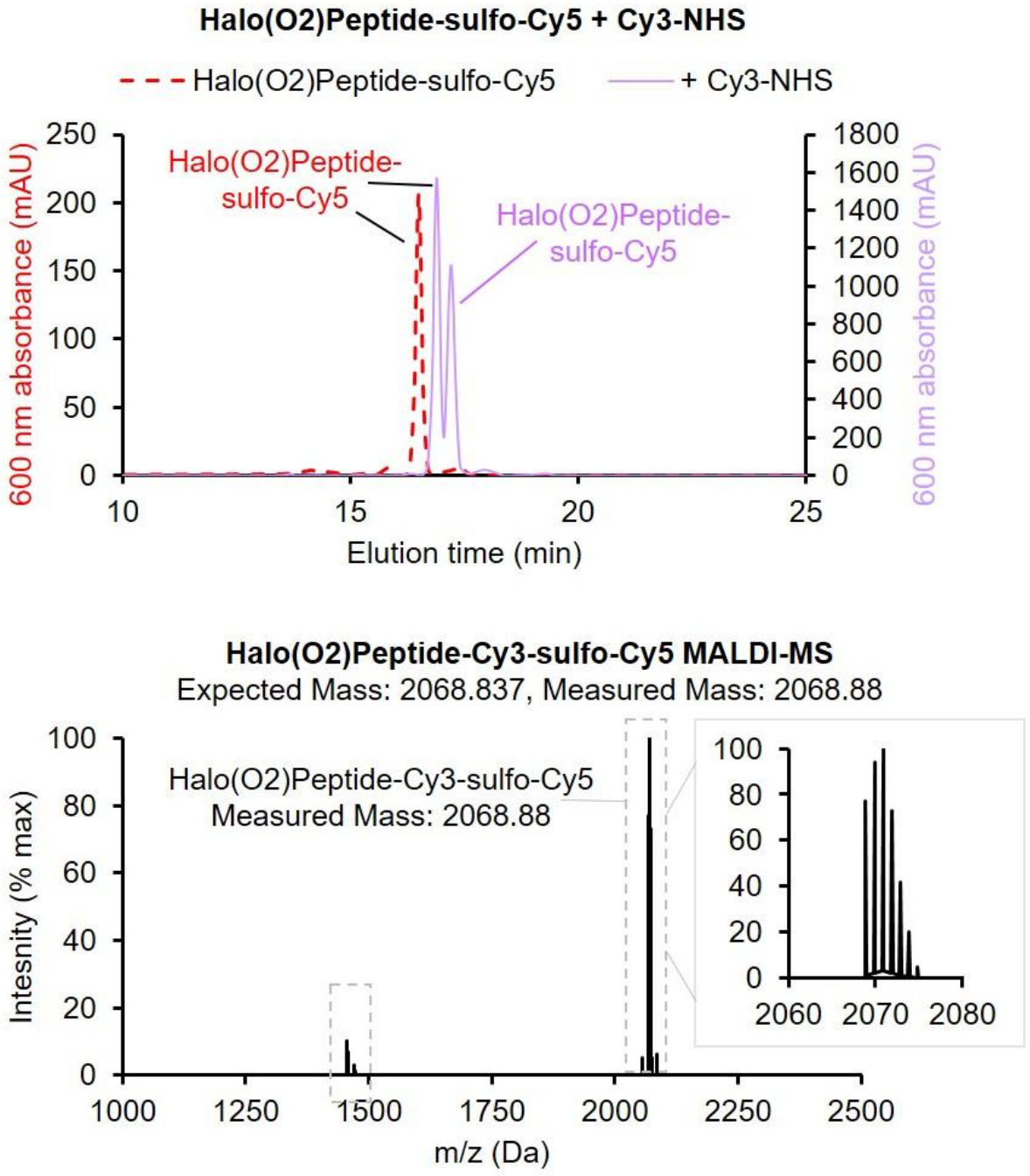


Figure 7.1 – Synthesis of Halo(O2)Peptide-Cy3-sulfo-Cy5. ***A*** *RP-HPLC trace of Halo(O2)Peptide-sulfo-Cy5 (synthesis shown in before and after reaction with Cy3-NHS).* ***B*** *MALDI-MS of the latter reaction peak from* ***A*** *has the expected mass of 2068.88 Da.*

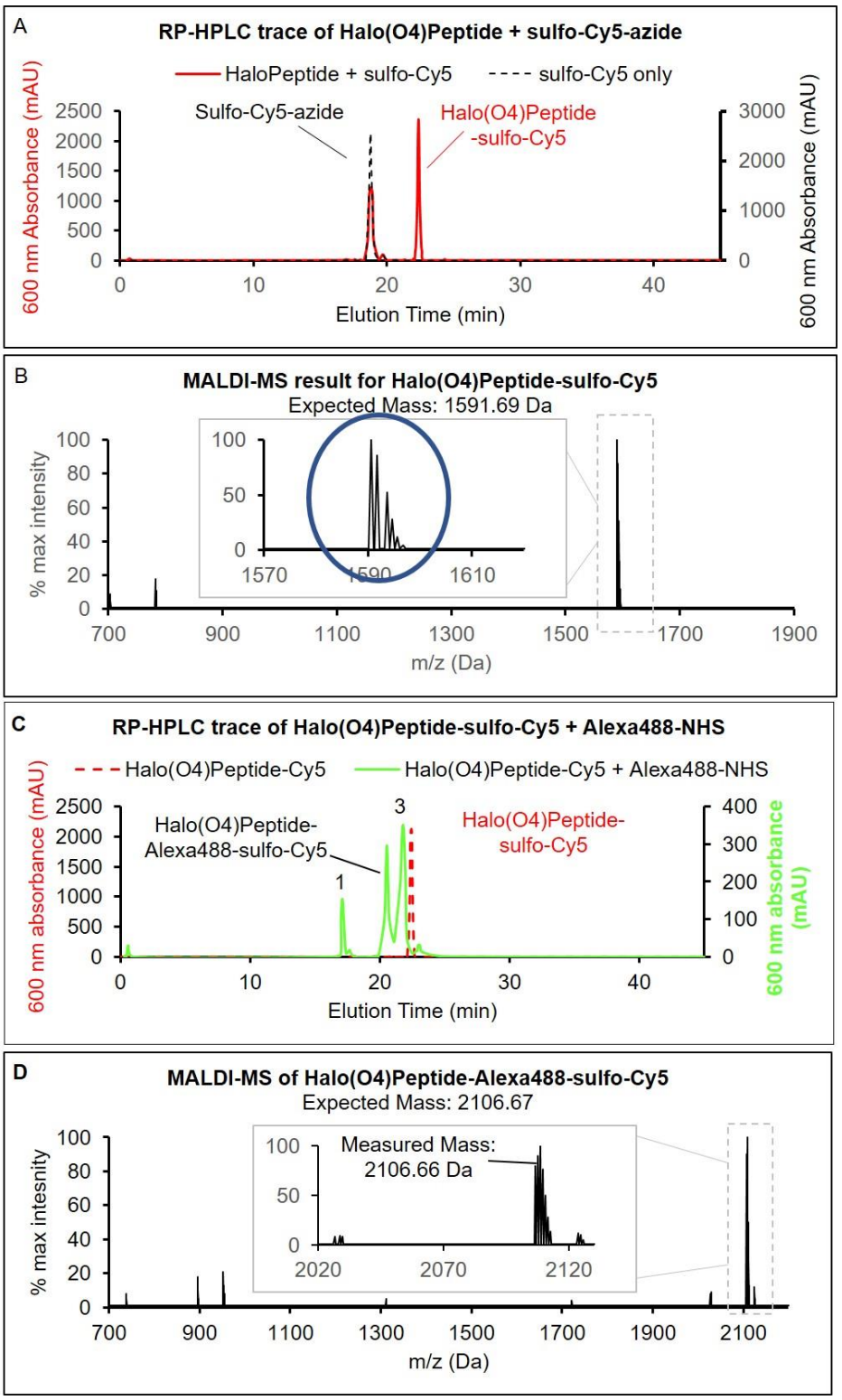


Figure 7.2 – Synthesis of Halo(O4)Peptide-Cy3-slufo-Cy5. A *HPLC separation of the products from the reaction in red, overlaid with the control reaction mixture containing only sulfo-Cy5, without Halo(O4)Peptide.* ***B*** *MALDI-MS verification that peak 2 from* ***A*** *is the intended product, with a measured mass of 1591.65 Da, equal to the predicated exact mass. In laid is a zoomed section on the major group of peaks. Each peak is separated from the previous by 1 Da.* ***C*** *HPLC separation of the products from the NHS reaction with the Halo(O4)Peptide-Cy5 shown in red dotted line, and the NHS reaction shown in green, multiple peaks are produced.* ***D*** *MALDI-MS verification that the peak 2 from* ***C*** *is the intended product. An inset on the main set of peaks is shown, and the first peak has a measured mass of 2106.66 Da, equal to the predicated exact mass.*

8 Bibliography

Ahmad, Zuhaida Asra, Swee Keong Yeap, Abdul Manaf Ali, Wan Yong Ho, Noorjahan Banu Mohamed Alitheen, and Muhajir Hamid. 2012. “ScFv Antibody: Principles and Clinical Application.” *Clinical and Developmental Immunology* 2012. doi:10.1155/2012/980250.

Aitken, Colin Echeverría, R. Andrew Marshall, and Joseph D. Puglisi. 2008. “An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments.” *Biophysical Journal* 94 (5): 1826–35. doi:10.1529/biophysj.107.117689.

Annibale, P., M. Scarselli, A. Kodiyan, and A. Radenovic. 2010. “Photoactivatable Fluorescent Protein MEos2 Displays Repeated Photoactivation after a Long-Lived Dark State in the Red Photoconverted Form.” *Journal of Physical Chemistry Letters* 1 (9): 1506–10. doi:10.1021/jz1003523.

Annibale, P., Marco Scarselli, Mattia Greco, and Aleksandra Radenovic. 2012. “Identification of the Factors Affecting Co-Localization Precision for Quantitative Multicolor Localization Microscopy.” *Optical Nanoscopy* 1 (1): 9. doi:10.1186/2192-2853-1-9.

Annibale, P., Stefano Vanni, Marco Scarselli, Ursula Rothlisberger, and Aleksandra Radenovic. 2011. “Quantitative Photo Activated Localization Microscopy: Unraveling the Effects of Photoblinking.” *PLoS ONE* 6 (7). doi:10.1371/journal.pone.0022678.

Aquino, Daniel, Andreas Schönle, Claudia Geisler, Claas V. Middendorff, Christian A. Wurm, Yosuke Okamura, Thorsten Lang, Stefan W. Hell, and Alexander Egner. 2011. “Two-Color Nanoscopy of Three-Dimensional Volumes by 4Pi Detection of Stochastically Switched Fluorophores.” *Nature Methods* 8 (4): 353–59. doi:10.1038/nmeth.1583.

Arbabi Ghahroudi, M., A. Desmyter, L. Wyns, R. Hamers, and S. Muyldermans. 1997. “Selection and Identification of Single Domain Antibody Fragments from Camel Heavy-Chain Antibodies.” *FEBS Letters* 414 (3). Federation of European Biochemical Societies: 521–26. doi:10.1016/S0014-5793(97)01062-4.

Baddeley, David, David Crossman, Sabrina Rossberger, Juliette E. Cheyne, Johanna

M. Montgomery, Isuru D. Jayasinghe, Christoph Cremer, Mark B. Cannell, and Christian Soeller. 2011. “4D Super-Resolution Microscopy with Conventional Fluorophores and Single Wavelength Excitation in Optically Thick Cells and Tissues.” *PLoS ONE* 6 (5). doi:10.1371/journal.pone.0020645.

Baeyer, A. von. 1871. “About a New Class of Dyes.” *Ber Deut Chem Ges* 4 (2). doi:10.1002/cber.18710040209.

Bartke, Till, Michiel Vermeulen, Blerta Xhemalce, Samuel C. Robson, Matthias Mann, and Tony Kouzarides. 2010. “Nucleosome-Interacting Proteins Regulated by DNA and Histone Methylation.” *Cell* 143 (3). Elsevier Inc.: 470–84. doi:10.1016/j.cell.2010.10.012.

Basu, Srinjan, Lisa Maria Needham, David Lando, Edward J.R. Taylor, Kai J. Wohlfahrt, Devina Shah, Wayne Boucher, et al. 2018. “FRET-Enhanced Photostability Allows Improved Single-Molecule Tracking of Proteins and Protein Complexes in Live Mammalian Cells.” *Nature Communications* 9 (1). doi:10.1038/s41467-018-04486-0.

Bates, M., Timothy R. Blosser, and Xiaowei Zhuang. 2005. “Short-Range Spectroscopic Ruler Based on a Single-Molecule Optical Switch.” *Physical Review Letters* 94 (10): 1–4. doi:10.1103/PhysRevLett.94.108101.

Bates, M., G. T. Dempsey, Kok Hao Chen, and Xiaowei Zhuang. 2012. “Multicolor Super-Resolution Fluorescence Imaging via Multi-Parameter Fluorophore Detection.” *ChemPhysChem* 13 (1): 99–107. doi:10.1002/cphc.201100735.

Bates, M., B. Huang, G. T. Dempsey, and X. Zhuang. 2007. “Multicolor Super-Resolution Imaging with Photo-Switchable Fluorescent Probes.” *Science* 317 (5845): 1749–53. doi:10.1126/science.1146598.

Beliveau, Brian J., Alistair N. Boettiger, Maier S. Avendaño, Ralf Jungmann, Ruth B. McCole, Eric F. Joyce, Caroline Kim-Kiselak, et al. 2015. “Single-Molecule Super-Resolution Imaging of Chromosomes and in Situ Haplotype Visualization Using Oligopaint FISH Probes.” *Nature Communications* 6 (May): 7147. doi:10.1038/ncomms8147.

Boder, Eric T., and K. Dane Wittrup. 1997. “Yeast Surface Display for Screening Combinatorial Polypeptide Libraries.” *Nature Biotechnology* 15 (6): 553–57.

doi:10.1038/nbt0697-553.

Brede, Norman, and M. Lakadamyali. 2012. “GraspJ: An Open Source, Real-Time Analysis Package for Super-Resolution Imaging.” *Optical Nanoscopy* 1 (1): 11. doi:10.1186/2192-2853-1-11.

Brewster, D. 1834. “On the Colours of Natural Bodies.” *Transactions of the Royal Society of Edinburgh* 12: 538–45.

Brodehl, Andreas, Per Niklas Hedde, Mareike Dieding, Azra Fatima, Volker Walhorn, Susan Gayda, Tomo Saric, et al. 2012. “Dual Color Photoactivation Localization Microscopy of Cardiomyopathy- Associated Desmin Mutants.” *Journal of Biological Chemistry* 287 (19): 16047–57. doi:10.1074/jbc.M111.313841.

Campbell, Robert E, Oded Tour, Amy E Palmer, Paul a Steinbach, Geoffrey S Baird, David a Zacharias, and Roger Y Tsien. 2002. “A Monomeric Red Fluorescent Protein.” *Proceedings of the National Academy of Sciences of the United States of America* 99 (12): 7877–82. doi:10.1073/pnas.082243699.

Chalfie, Martin, Yuan Tu, Ghia Euskirchen, William W Ward, and Douglas C Prashert. 1994. “Green Fluorescent Protein as a Marker for Gene Expression Author (s): Martin Chalfie , Yuan Tu , Ghia Euskirchen , William W . Ward , Douglas C . Prasher Published by : American Association for the Advancement of Science Stable URL : Http://Www.Jstor.Or.” *Science* 263 (5148): 802–5.

Chaumeil, J, Patricia Le Baccon, Anton Wutz, and E Heard. 2006. “A Novel Role for Xist RNA in the Formation of a Repressive Nuclear Compartment into Which Genes Are Recruited When Silenced.” *Genes and Development* 20 (16): 2223–37. doi:10.1101/gad.380906.

Chaumeil, J, J Chow, S Augui, and E Heard. 2008. “Combined Immunofluorescence, RNA Fluorescent in Situ Hybridization to Study Chromatin Changes, Transcriptional Activity, Nuclear Organization” 1: 1–9.

Choo, Andre B H, Rosanne D Dunn, Kevin W Broady, and Robert L Raison. 2002. “Soluble Expression of a Functional Recombinant Cytolytic Immunotoxin in Insect Cells.” *Protein Expression and Purification* 24 (3): 338–47. doi:10.1006/prep.2001.1589.

Choo, Khar Heng, Joo Chuan Tong, and Shoba Ranganathan. 2007. “Modeling Escherichia Coli Signal Peptidase Complex with Bound Substrate: Determinants in the Mature Peptide Influencing Signal Peptide Cleavage.” *BMC Bioinformatics* 9 (9): 27–30. doi:10.1186/1471-2105-9-S1-S15.

Clackson, Tim, Hennie R. Hoogenboom, Andrew D. Griffiths, and Greg Winter. 1991. “Making Antibody Fragments Using Phage Display Libraries.” *Nature* 352 (6336): 624–28. doi:10.1038/352624a0.

Cognet, Laurent, and Brahim Lounis. 2014. “Advances in Live-Cell Single-Particle Tracking and Dynamic Super-Resolution Imaging To Cite This Version :,” 78–85.

Coltharp, Carla, Rene P. Kessler, and Jie Xiao. 2012. “Accurate Construction of Photoactivated Localization Microscopy (PALM) Images for Quantitative Measurements.” *PLoS ONE* 7 (12). doi:10.1371/journal.pone.0051725.

Conley, Nicholas R., Julie S. Biteen, and W. E. Moerner. 2008. “Cy3-Cy5 Covalent Heterodimers for Single-Molecule Photoswitching.” *Journal of Physical Chemistry B* 112 (38): 11878–80. doi:10.1021/jp806698p.

Coons, A. H., H. J. Creech, and N. R. Jones. 1941. “Immunological Properties of an Antibody Conatining a Fluorescent Group.” *Proceedings for the Society of Experimental Biology and Medicine* 47 (2): 200–202. doi:10.3181/00379727-47-13084P.

Cranfill, Paula J, Brittney R Sell, Michelle A Baird, John R Allen, Zeno Lavagnino, H Martijn de Gruiter, Gert-Jan Kremers, Michael W Davidson, Alessandro Ustione, and David W Piston. 2016. “Quantitative Assessment of Fluorescent Proteins.” *Nature Methods* 13 (7): 557–62. doi:10.1038/nmeth.3891.

Cubitt, a B, Roger Heim, Stephen R Adams, a E Boyd, Larry a Gross, and Roger Y Tsien. 1995. “Understanding, Improving and Using Green Fluorescent Proteins.” *Trends in Biochemical Sciences* 20 (11): 448–55. doi:10.1016/S0968-0004(00)89099-4.

Dawson, P E, Churchill, M J, Ghadiri, M R, Kent S B H, Philip E. Dawson, Michael J. Churchill, M. Reza Ghadiri, and Stephen B.H. Kent. 1997. “Modulation Fo Reactivity in Native Chemical Ligation through the Use of Thiol Additive.” *Journal of the American Chemical Society* 119 (19): 4325–29.

Dawson, P., T. Muir, I Clark-Lewis, and S. Kent. 1994. “Synthesis of Proteins by Native Chemical Ligation.” *Science* 266 (5186): 776–79. doi:10.1126/science.7973629.

Dellagiacoma, C., G. Lukinavičius, N. Bocchio, S. Banala, S. Geissbühler, I. Märki, K. Johnsson, and T. Lasser. 2010. “Targeted Photoswitchable Probe for Nanoscopy of Biological Structures.” *ChemBioChem* 11 (10): 1361–63. doi:10.1002/cbic.201000189.

Dempsey, G. T., M. Bates, Walter E. Kowtoniuk, David R. Liu, Roger Y. Tsien, and Xiaowei Zhuang. 2009. “Photoswitching Mechanism of Cyanine Dyes.” *Journal of the American Chemical Society* 131 (51): 18192–93. doi:10.1021/ja904588g.

Dempsey, G. T., Joshua C Vaughan, Kok Hao Chen, M. Bates, and Xiaowei Zhuang. 2011. “Evaluation of Fluorophores for Optimal Performance in Localization-Based Super-Resolution Imaging.” *Nature Methods* 8 (12): 1027–36. doi:10.1038/nmeth.1768.

Devaraj, Neal K, and Ralph Weissleder. 2011. “Biomedical Applications of Tetrazine Cycloadditions.” *Society* 44 (9): 816–27. doi:10.1021/ar200037t.

Dries, K. van den, S. L. Schwartz, J. Byars, M. B. M. Meddens, M. Bolomini-Vittori, D. S. Lidke, C. G. Figdor, K. A. Lidke, and A. Cambi. 2013. “Dual-Color Superresolution Microscopy Reveals Nanoscale Organization of Mechanosensory Podosomes.” *Molecular Biology of the Cell* 24 (13): 2112–23. doi:10.1091/mbc.E12-12-0856.

Durisic, N., Lara Laparra Cuervo, and M. Lakadamyali. 2014. “Quantitative Super-Resolution Microscopy: Pitfalls and Strategies for Image Analysis.” *Current Opinion in Chemical Biology* 20 (1). Elsevier Ltd: 22–28. doi:10.1016/j.cbpa.2014.04.005.

Durisic, N., A. G. Godin, C. M. Wever, C. D. Heyes, M. Lakadamyali, and J. A. Dent. 2012. “Stoichiometry of the Human Glycine Receptor Revealed by Direct Subunit Counting.” *Journal of Neuroscience* 32 (37): 12915–20. doi:10.1523/JNEUROSCI.2050-12.2012.

Edelstein, Arthur, Nenad Amodaj, Karl Hoover, Ron Vale, and Nico Stuurman. 2010. “Computer Control of Microscopes Using Manager.” *Current Protocols in Molecular Biology*, no. SUPPL. 92: 1–17. doi:10.1002/0471142727.mb1420s92.

Edelstein, Arthur D, Mark A Tsuchida, Nenad Amodaj, Henry Pinkard, Ronald D Vale, and Nico Stuurman. 2014. “Advanced Methods of Microscope Control Using MManager Software.” *Journal of Biological Methods* 1 (2): 10. doi:10.14440/jbm.2014.36.

Flajnik, Martin F, Nick Deschacht, and Serge Muyldermans. 2011. “A Case Of Convergence : Why Did a Simple Alternative to Canonical Antibodies Arise in Sharks and Camels ?” 9 (8). doi:10.1371/journal.pbio.1001120.

Fornasiero, Eugenio F., and Felipe Opazo. 2015. “Super-Resolution Imaging for Cell Biologists.” *BioEssays* 37 (4): 436–51. doi:10.1002/bies.201400170.

Franken, Sybille M, Henriette J Rozeboom, Kor H Kalk, Bauke W Dijkstra, and W G J Hol. 1991. “Crystal Structure of Haloalkane Dehalogenase: An Enzyme to Detoxify Halogenated Alkanes.” *The EMBO Journal* 10 (6): 1297–1302.

Fridy, Peter C, Yinyin Li, Sarah Keegan, Mary K Thompson, Ilona Nudelman, Johannes F Scheid, Marlene Oeffinger, et al. 2014. “A Robust Pipeline for Rapid Production of Versatile Nanobody Repertoires.” *Nature Methods* 11 (12): 1253–60. doi:10.1038/nmeth.3170.

Frost, Nicholas A., Hari Shroff, Huihui Kong, E. Betzig, and Thomas A. Blanpied. 2010. “Single-Molecule Discrimination of Discrete Perisynaptic and Distributed Sites of Actin Filament Assembly within Dendritic Spines.” *Neuron* 67 (1). Elsevier Ltd: 86–99. doi:10.1016/j.neuron.2010.05.026.

Gahlmann, Andreas, and W. E. Moerner. 2013. “Exploring Bacterial Cell Biology with Single-Molecule Tracking and Super-Resolution Imaging.” *Nature Reviews Microbiology* 12 (1): 9–22. doi:10.1038/nrmicro3154.

Gao, Jing, Ye Wang, Mingjun Cai, Yangang Pan, Haijiao Xu, Junguang Jiang, Hongbin Ji, and Hongda Wang. 2015. “Mechanistic Insights into EGFR Membrane Clustering Revealed by Super-Resolution Imaging.” *Nanoscale* 7 (6). Royal Society of Chemistry: 2511–19. doi:10.1039/C4NR04962D.

Giannone, Gregory, Eric Hosy, Florian Levet, Audrey Constals, Katrin Schulze, Alexander I. Sobolevsky, Michael P. Rosconi, et al. 2010. “Dynamic Superresolution Imaging of Endogenous Proteins on Living Cells at Ultra-High Density.” *Biophysical Journal* 99 (4): 1303–10. doi:10.1016/j.bpj.2010.06.005.

Greenberg, Andrew S;, David Avila, Marianne Hughes, Austin Hughes, E. Churchill McKinney, and Martin F. Flajnik. 1995. “A New Antigen Receptor Gene Family That Undergoes Rearrangement And ...”

Grimm, Jonathan B, Brian P English, Jiji Chen, Joel P Slaughter, Zhengjian Zhang, Andrey Revyakin, Ronak Patel, et al. 2015. “A General Method to Improve Fluorophores for Live-Cell and Single-Molecule Microscopy.” *Nature Methods* 12 (3): 244–50. doi:10.1038/nmeth.3256.

Grimm, Jonathan B, Brian P English, Heejun Choi, Anand K Muthusamy, Brian P Mehl, Peng Dong, Timothy A Brown, et al. 2016. “Bright Photoactivatable Fluorophores for Single-Molecule Imaging.” *Nature Methods* 13 (12): 985–88. doi:10.1038/nmeth.4034.

Grunwaldt, Gisela. 2002. “Grunwaldt et Al. - 2002 - M Ultiple Binding Sites of Uorescein Isothiocyanate Moieties on Myoglobin Photophysical Heterogeneity as Revealed by Ground- and Excited-State Spectroscopy.Pdf” 67: 177–86.

Habuchi, Satoshi, Peter Dedecker, Jun-ichi Hotta, Cristina Flors, Ryoko Ando, Hideaki Mizuno, Atsushi Miyawaki, and Johan Hofkens. 2006. “Photo-Induced Protonation/Deprotonation in the GFP-like Fluorescent Protein Dronpa: Mechanism Responsible for the Reversible Photoswitching.” *Photochem. Photobiol. Sci.* 5: 567–76. doi:10.1039/b516339k.

Hamers-Casterman, C, T. Atarhouch, S. Muyldermans, G Robinson, C. Hammers, E. Bajyana Song, N. Bendahman, and R. Hammers. 1993. “Naturally Occurring Antibodies Devoid of Light Chains.” *Nature* 363: 446–48. doi:10.1038/363446a0.

Han, Chen, and Jianping Wang. 2012. “Influence of an Unnatural Amino Acid Side Chain on the Conformational Dynamics of Peptides.” *ChemPhysChem* 13 (6): 1522–34. doi:10.1002/cphc.201100995.

Hanly, W. C., J. E. Artwohl, and B. T. Bennett. 1995. “Review of Polyclonal Antibody Production Procedures in Mammals and Poultry.” *ILAR Journal* 37 (3): 93–118. doi:10.1093/ilar.37.3.93.

Hauser, Paul S., and Robert O. Ryan. 2007. “Expressed Protein Ligation Using an N-Terminal Cysteine Containing Fragment Generated in Vivo from a PelB Fusion Protein.” *Protein Expression and Purification* 54 (2): 227–33.

doi:10.1016/j.pep.2007.04.002.

Heilemann, M., S. Van De Linde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. 2008. “Subdiffraction-Resolution Fluorescence Imaging with Conventional Fluorescent Probes.” *Angewandte Chemie - International Edition* 47 (33): 6172–76. doi:10.1002/anie.200802376.

Heilemann, M., Emmanuel Margeat, Robert Kasper, Markus Sauer, and Philip Tinnefeld. 2005. “Carbocyanine Dyes as Efficient Reversible Single-Molecule Optical Switch.” *Journal of the American Chemical Society* 127 (11): 3801–6. doi:10.1021/ja044686x.

Herbert, Alex. 2014. “PeakFit, a Single-Molecule Plugins.” http://www.sussex.ac.uk/gdsc/intranet/microscopy/imagej/smlm_plugins.

Herschel, J. F. T. W. 1845. “On a Case of Superficial Colour Presented by a Homogeneous Liquid Internally Colourless.” *Philosophical Transactions of the Royal Society of London* 135: 143–45. http://www.jstor.org/stable/108266.

Hsu, Chih Jung, and Tobias Baumgart. 2011. “Spatial Association of Signaling Proteins and F-Actin Effects on Cluster Assembly Analyzed via Photoactivation Localization Microscopy in T Cells.” *PLoS ONE* 6 (8). doi:10.1371/journal.pone.0023586.

Huang, B., Sara A Jones, Boerries Brandenburg, and Xiaowei Zhuang. 2008. “Whole-Cell 3D STORM Reveals Interactions between Cellular Structures with Nanometer-Scale Resolution.” *Nature Methods* 5 (12): 1047–52. doi:10.1038/nmeth.1274.

Huang, B., W. Wang, M. Bates, and X. Zhuang. 2008. “Three-Dimensional Super-Resolution Imaging by Stochastic Optical Reconstruction Microscopy.” *Science* 319 (5864): 810–13. doi:10.1126/science.1153529.

Hummer, G., F. Fricke, and M. Heilemann. 2016. “Model-Independent Counting of Molecules in Single-Molecule Localization Microscopy.” *Molecular Biology of the Cell* 27 (22): 3637–44. doi:10.1091/mbc.E16-07-0525.

Inouye, Satoshi, and Frederick I. Tsuji. 1994. “Aequorea Green Fluorescent Protein.”

FEBS Letters 341 (2–3): 277–80. doi:10.1016/0014-5793(94)80472-9.

Izenman, Alan Julian. 1991. “Review Papers: Recent Developments in Nonparametric Density Estimation.” *Journal of the American Statistical Association* 86 (413). Taylor & Francis: 205–24. doi:10.1080/01621459.1991.10475021.

Jones, Sara A, Sang-Hee Shim, Jiang He, and Xiaowei Zhuang. 2011. “Fast, Three-Dimensional Super-Resolution Imaging of Live Cells.” *Nature Methods* 8 (6): 499–505. doi:10.1038/nmeth.1605.

Jung, S, and A. Pluckthun. 1997. “Improving in Vivo Folding and Stability of a Single-Chain Fv Antibody Fragment by Loop Grafting.” *Protein Engineering Design and Selection* 10 (8): 959–66. doi:10.1093/protein/10.8.959.

Jungmann, Ralf, Maier S Avendaño, Mingjie Dai, Johannes B Woehrstein, Sarit S Agasti, Zachary Feiger, Avital Rodal, and Peng Yin. 2016. “Quantitative Super-Resolution Imaging with QPAINT.” *Nature Methods* 13 (5): 439–42. doi:10.1038/nmeth.3804.

Jungmann, Ralf, Maier S Avendaño, Johannes B Woehrstein, Mingjie Dai, William M Shih, and Peng Yin. 2014. “Multiplexed 3D Cellular Super-Resolution Imaging with DNA-PAINT and Exchange-PAINT.” *Nature Methods* 11 (3): 313–18. doi:10.1038/nmeth.2835.

Jungmann, Ralf, Christian Steinhauer, Max Scheible, Anton Kuzyk, Philip Tinnefeld, and Friedrich C. Simmel. 2010. “Single-Molecule Kinetics and Super-Resolution Microscopy by Fluorescence Imaging of Transient Binding on DNA Origami.” *Nano Letters* 10 (11): 4756–61. doi:10.1021/nl103427w.

Karathanasis, C., F. Fricke, G. Hummer, and M. Heilemann. 2017. “Molecule Counts in Localization Microscopy with Organic Fluorophores.” *ChemPhysChem* 18 (8): 942–48. doi:10.1002/cphc.201601425.

Kaufmann, Markus, Peter Lindner, Annemarie Honegger, Kerstin Blank, Markus Tschopp, Guido Capitani, Andreas Plückthun, and Markus G. Grütter. 2002. “Crystal Structure of the Anti-His Tag Antibody 3D5 Single-Chain Fragment Complexed to Its Antigen.” *Journal of Molecular Biology* 318 (1): 135–47. doi:10.1016/S0022-2836(02)00038-4.

Kirchhofer, Axel, Jonas Helma, Katrin Schmidthals, Carina Frauer, Sheng Cui, Annette Karcher, Mireille Pellis, et al. 2010. “Modulation of Protein Properties in Living Cells Using Nanobodies.” *Nature Structural & Molecular Biology* 17 (1). Nature Publishing Group: 133–38. doi:10.1038/nsmb.1727.

Klein, Teresa, Anna Löschberger, Sven Proppert, S. Wolter, S. Van De Linde, and Markus Sauer. 2011. “Live-Cell DSTORM with SNAP-Tag Fusion Proteins.” *Nature Methods* 8 (1): 7–9. doi:10.1038/nmeth0111-7b.

KÖHLER, G., and C. MILSTEIN. 1975. “Continuous Cultures of Fused Cells Secreting Antibody of Predefined Specificity.” *Nature* 256 (5517): 495–97. doi:10.1038/256495a0.

Kolb, Hartmuth C, M G Finn, and K Barry Sharpless. 2001. “Click Chemistry: Diveverse Chemical Funtion from a Few Good Reactions.Pdf.” *Angewandte Chemie - International Edition* 40: 2004–21. doi:10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.3.CO;2-X.

Kosaka, Nobuyuki, Mikako Ogawa, Peter L. Choyke, Natasha Karassina, Cesear Corona, Mark McDougall, David T. Lynch, et al. 2009. “In Vivo Stable Tumor-Specific Painting in Various Colors Using Dehalogenase-Based Protein-Tag Fluorescent Ligands.” *Bioconjugate Chemistry* 20 (7): 1367–74. doi:10.1021/bc9001344.

Krauß, Norbert, Helga Wessner, Karin Welfle, Heinz Welfle, Christa Scholz, Martina Seifert, Kristina Zubow, et al. 2008. “The Structure of the Anti-c-Myc Antibody 9E10 Fab Fragment/Epitope Peptide Complex Reveals a Novel Binding Mode Dominated by the Heavy Chain Hypervariable Loops.” *Proteins: Structure, Function and Genetics* 73 (3): 552–65. doi:10.1002/prot.22080.

Krüger, Carmen, Franziska Fricke, Christos Karathanasis, Marina S. Dietz, Sebastian Malkusch, Gerhard Hummer, and Mike Heilemann. 2017. “Molecular Counting of Membrane Receptor Subunits with Single-Molecule Localization Microscopy,” no. February 2017: 100710K. doi:10.1117/12.2256638.

Kubala, Marta H., Oleksiy Kovtun, Kirill Alexandrov, and Brett M. Collins. 2010. “Structural and Thermodynamic Analysis of the GFP:GFP-Nanobody Complex.” *Protein Science* 19 (12): 2389–2401. doi:10.1002/pro.519.

Lakadamyali, M., Hazen Babcock, M. Bates, Xiaowei Zhuang, and Jeff Lichtman. 2012. “3D Multicolor Super-Resolution Imaging Offers Improved Accuracy in Neuron Tracing.” *PLoS ONE* 7 (1). doi:10.1371/journal.pone.0030826.

Lampe, André, Volker Haucke, Stephan J. Sigrist, M. Heilemann, and Jan Schmoranzer. 2012. “Multi-Colour Direct STORM with Red Emitting Carbocyanines.” *Biology of the Cell* 104 (4): 229–37. doi:10.1111/boc.201100011.

Lando, D., U. Endesfelder, H. Berger, L. Subramanian, P. D. Dunne, J. McColl, D. Klenerman, et al. 2012. “Quantitative Single-Molecule Microscopy Reveals That CENP-ACnp1 Deposition Occurs during G2 in Fission Yeast.” *Open Biology* 2 (7): 120078–120078. doi:10.1098/rsob.120078.

Leake, Mark C., Jennifer H. Chandler, George H. Wadhams, Fan Bai, Richard M. Berry, and Judith P. Armitage. 2006. “Stoichiometry and Turnover in Single, Functioning Membrane Protein Complexes.” *Nature* 443 (7109): 355–58. doi:10.1038/nature05135.

Lee, S.-H., J. Y. Shin, A. Lee, and C. Bustamante. 2012. “Counting Single Photoactivatable Fluorescent Molecules by Photoactivated Localization Microscopy (PALM).” *Proceedings of the National Academy of Sciences* 109 (43): 17436–41. doi:10.1073/pnas.1215175109.

Legant, Wesley R, Lin Shao, Jonathan B Grimm, Timothy A Brown, Daniel E Milkie, Brian B Avants, Luke D Lavis, and Eric Betzig. 2016. “High-Density Three-Dimensional Localization Microscopy across Large Volumes.” *Nature Methods* 13 (4): 359–65. doi:10.1038/nmeth.3797.

Leinfelder, W, K Forchhammer, B Veprek, E Zehelein, and a Böck. 1990. “In Vitro Synthesis of Selenocysteinyl-TRNA(UCA) from Seryl-TRNA(UCA): Involvement and Characterization of the SelD Gene Product.” *Proceedings of the National Academy of Sciences of the United States of America* 87 (January): 543–47. doi:10.1073/pnas.87.2.543.

Lew, M. D., S. F. Lee, J. L. Ptacin, M. K. Lee, R. J. Twieg, L. Shapiro, and W. E. Moerner. 2011. “Three-Dimensional Superresolution Colocalization of Intracellular Protein Superstructures and the Cell Surface in Live Caulobacter Crescentus.” *Proceedings of the National Academy of Sciences* 108 (46): E1102–10.

doi:10.1073/pnas.1114444108.

Linde, S. Van De, Sarah Aufmkolk, Christian Franke, Thorge Holm, Teresa Klein, A. Löschberger, Sven Proppert, S. Wolter, and Markus Sauer. 2013. “Investigating Cellular Structures at the Nanoscale with Organic Fluorophores.” *Chemistry and Biology* 20 (1): 8–18. doi:10.1016/j.chembiol.2012.11.004.

Linde, S. Van De, R. Kasper, M. Heilemann, and M. Sauer. 2008. “Photoswitching Microscopy with Standard Fluorophores.” *Applied Physics B: Lasers and Optics* 93 (4): 725–31. doi:10.1007/s00340-008-3250-9.

Linde, S. Van De, Ivan Krstić, Thomas Prisner, Sören Doose, M. Heilemann, and Markus Sauer. 2011. “Photoinduced Formation of Reversible Dye Radicals and Their Impact on Super-Resolution Imaging.” *Photochem. Photobiol. Sci.* 10 (4): 499–506. doi:10.1039/C0PP00317D.

Linde, Sebastian Van De, Anna Löschberger, Teresa Klein, Meike Heidbreder, Steve Wolter, Mike Heilemann, and Markus Sauer. 2011. “Direct Stochastic Optical Reconstruction Microscopy with Standard Fluorescent Probes.” *Nature Protocols* 6 (7): 991–1009. doi:10.1038/nprot.2011.336.

Linden, R H van Der, L G Frenken, B de Geus, M M Harmsen, R C Ruuls, W Stok, L de Ron, S Wilson, P Davis, and C T Verrips. 1999. “Comparison of Physical Chemical Properties of Llama VHH Antibody Fragments and Mouse Monoclonal Antibodies.” *Biochimica et Biophysica Acta* 1431 (1): 37–46. doi:10.1016/S0167-4838(99)00030-8.

Liu, Yu, Matthew Fares, Noah P. Dunham, Zi Gao, Kun Miao, Xueyuan Jiang, Samuel S. Bollinger, Amie K. Boal, and Xin Zhang. 2017. “AgHalo: A Facile Fluorogenic Sensor to Detect Drug-Induced Proteome Stress.” *Angewandte Chemie - International Edition* 56 (30): 8672–76. doi:10.1002/anie.201702417.

Lobstein, Julie, Charlie A Emrich, Chris Jeans, Melinda Faulkner, Paul Riggs, and Mehmet Berkmen. 2012. “SHuffle, a Novel Escherichia Coli Protein Expression Strain Capable of Correctly Folding Disulfide Bonded Proteins in Its Cytoplasm.” *Microbial Cell Factories* 11 (1). BioMed Central Ltd: 753. doi:10.1186/1475-2859-11-56.

Los, Georgyi V, Lance P Encell, Mark G McDougall, Danette D Hartzell, Natasha

Karassina, Chad Zimprich, Monika G Wood, et al. 2008. “HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis.” *ACS Chemical Biology* 3 (6). American Chemical Society: 373–82. doi:10.1021/cb800025k.

Löschberger, A., S. van de Linde, M.-C. Dabauvalle, B. Rieger, M. Heilemann, G. Krohne, and M. Sauer. 2012. “Super-Resolution Imaging Visualizes the Eightfold Symmetry of Gp210 Proteins around the Nuclear Pore Complex and Resolves the Central Channel with Nanometer Resolution.” *Journal of Cell Science* 125 (3): 570–75. doi:10.1242/jcs.098822.

Löschberger, Anna, Thomas Niehörster, and Markus Sauer. 2014. “Click Chemistry for the Conservation of Cellular Structures and Fluorescent Proteins: ClickOx.” *Biotechnology Journal* 9 (5): 693–97. doi:10.1002/biot.201400026.

Lubeck, Eric, and Long Cai. 2012. “Single-Cell Systems Biology by Super-Resolution Imaging and Combinatorial Labeling.” *Nature Methods* 9 (7): 743–48. doi:10.1038/nmeth.2069.

Lukinavičius, Gražvydas, Keitaro Umezawa, Nicolas Olivier, Alf Honigmann, Guoying Yang, Tilman Plass, Veronika Mueller, et al. 2013. “A Near-Infrared Fluorophore for Live-Cell Super-Resolution Microscopy of Cellular Proteins.” *Nature Chemistry* 5 (2): 132–39. doi:10.1038/nchem.1546.

Malkusch, Sebastian, Ulrike Endesfelder, Justine Mondry, Márton Gelléri, Peter J. Verveer, and M. Heilemann. 2012. “Coordinate-Based Colocalization Analysis of Single-Molecule Localization Microscopy Data.” *Histochemistry and Cell Biology* 137 (1): 1–10. doi:10.1007/s00418-011-0880-5.

Malkusch, Sebastian, and Mike Heilemann. 2016. “Extracting Quantitative Information from Single-Molecule Super-Resolution Imaging Data with LAMA - LocAlization Microscopy Analyzer.” *Scientific Reports* 6 (September). Nature Publishing Group: 6–9. doi:10.1038/srep34486.

Marquardt, D. 1963. “An Algorithm for Least-Squares Estimation of Nonlinear Parameters.” *Journal of the Society for Industrial and Applied Mathematics* 11 (2). Society for Industrial and Applied Mathematics: 431–41. doi:10.1137/0111030.

Martin, Cecile D, Gertrudis Rojas, Joanne N Mitchell, Karen J Vincent, Jiahua Wu, John McCafferty, and Darren J Schofield. 2006. “A Simple Vector System to

Improve Performance and Utilisation of Recombinant Antibodies." *BMC Biotechnology* 6: 46. doi:10.1186/1472-6750-6-46.

McKinney, Sean A, Christopher S Murphy, Kristin L Hazelwood, Michael W Davidson, and Loren L Looger. 2009. "A Bright and Photostable Photoconvertible Fluorescent Protein." *Nature Methods* 6 (2): 131–33. doi:10.1038/nmeth.1296.

Miller, Brian R., Stephen J. Demarest, Alexey Lugovskoy, Flora Huang, Xiufeng Wu, William B. Snyder, Lisa J. Croner, et al. 2010. "Stability Engineering of ScFvs for the Development of Bispecific and Multivalent Antibodies." *Protein Engineering, Design and Selection* 23 (7): 549–57. doi:10.1093/protein/gzq028.

Moon, Seonah, Rui Yan, Samuel J. Kenny, Yennie Shyu, Limin Xiang, Wan Li, and Ke Xu. 2017. "Spectrally Resolved, Functional Super-Resolution Microscopy Reveals Nanoscale Compositional Heterogeneity in Live-Cell Membranes." *Journal of the American Chemical Society* 139 (32): 10944–47. doi:10.1021/jacs.7b03846.

Muranyi, Walter, Sebastian Malkusch, Barbara Muller, M. Heilemann, and Hans Georg Krausslich. 2013. "Super-Resolution Microscopy Reveals Specific Recruitment of HIV-1 Envelope Proteins to Viral Assembly Sites Dependent on the Envelope C-Terminal Tail." *PLoS Pathogens* 9 (2). doi:10.1371/journal.ppat.1003198.

Nanguneri, Siddharth, Benjamin Flottmann, Heinz Horstmann, M. Heilemann, and Thomas Kuner. 2012. "Three-Dimensional, Tomographic Super-Resolution Fluorescence Imaging of Serially Sectioned Thick Samples." *PLoS ONE* 7 (5): 1–8. doi:10.1371/journal.pone.0038098.

Newman, Janet, Thomas S. Peat, Ruth Richard, Lynn Kan, Paul E. Swanson, Joseph A. Affholter, Ian H. Holmes, John F. Schindler, Clifford J. Unkefer, and Thomas C. Terwilliger. 1999. "Haloalkane Dehalogenases: Structure of a Rhodococcus Enzyme." *Biochemistry* 38 (49): 16105–14. doi:10.1021/bi9913855.

Nikić, Ivana, Jun Hee Kang, Gemma Estrada Girona, Iker Valle Aramburu, and Edward A. Lemke. 2015. "Labeling Proteins on Live Mammalian Cells Using Click Chemistry." *Nature Protocols* 10 (5): 780–91. doi:10.1038/nprot.2015.045.

Nikić, Ivana, and Edward A. Lemke. 2015. "Genetic Code Expansion Enabled Site-Specific Dual-Color Protein Labeling: Superresolution Microscopy and Beyond." *Current Opinion in Chemical Biology* 28: 164–73. doi:10.1016/j.cbpa.2015.07.021.

Olivier, Nicolas, Debora Keller, Pierre Gönczy, and Suliana Manley. 2013. “Resolution Doubling in 3D-STORM Imaging through Improved Buffers.” *PLoS ONE* 8 (7): 1–9. doi:10.1371/journal.pone.0069004.

Ollis, David L, Eong Cheah, Miroslaw Cygler, Bauke Dijkstra, Felix Frolow, Sybille M Franken, Michal Harel, et al. 1992. “The a / | 3 Hydrolase Fold” 5 (3): 197–211.

Owen, Dylan M., Carles Rentero, Jérémie Rossy, Astrid Magenau, David Williamson, Macarena Rodriguez, and Katharina Gaus. 2010. “PALM Imaging and Cluster Analysis of Protein Heterogeneity at the Cell Surface.” *Journal of Biophotonics* 3 (7): 446–54. doi:10.1002/jbio.200900089.

Palayret, Matthieu, Helen Armes, Srinjan Basu, Adam T. Watson, Alex Herbert, David Lando, Thomas J. Etheridge, et al. 2015. “Virtual-’light-Sheet’ Single-Molecule Localisation Microscopy Enables Quantitative Optical Sectioning for Super-Resolution Imaging.” *PLoS ONE* 10 (4): 1–15. doi:10.1371/journal.pone.0125438.

Pereira, Cândida F, Jérémie Rossy, Dylan M Owen, Johnson Mak, and Katharina Gaus. 2012. “HIV Taken by STORM: Super-Resolution Fluorescence Microscopy of a Viral Infection.” *Virology Journal* 9: 84. doi:10.1186/1743-422X-9-84.

Pinotsi, Dorothea, Alexander Kai Büll, Céline Galvagnion, M Dobson, Gabriele S Kaminski Schierle, and Clemens F Kaminski. 2013. “Kinetics via Two-Color Super-Resolution Microscopy Direct Observation of Heterogeneous Amyloid Fibril Growth Kinetics via Two-Color Super-Resolution Microscopy.”

Plass, Tilman, Sigrid Milles, Christine Koehler, Carsten Schultz, and Edward A. Lemke. 2011. “Genetically Encoded Copper-Free Click Chemistry.” *Angewandte Chemie - International Edition* 50 (17): 3878–81. doi:10.1002/anie.201008178.

Plass, Tilman, Sigrid Milles, Christine Koehler, Jędrzej Szymański, Rainer Mueller, Manfred Wießler, Carsten Schultz, and Edward A. Lemke. 2012. “Amino Acids for Diels-Alder Reactions in Living Cells.” *Angewandte Chemie - International Edition* 51 (17): 4166–70. doi:10.1002/anie.201108231.

Pleiner, Tino, M. Bates, Sergei Trakhanov, Chung Tien Lee, Jan Erik Schliep, Hema Chug, Marc Böhning, Holger Stark, Henning Urlaub, and Dirk Görlich. 2015. “Nanobodies: Site-Specific Labeling for Super-Resolution Imaging, Rapid Epitope- Mapping and Native Protein Complex Isolation.” *ELife* 4

(DECEMBER2015): 1–21. doi:10.7554/eLife.11349.

Prasher, Douglas C., Virginia K. Eckenrode, William W. Ward, Frank G. Prendergast, and Milton J. Cormier. 1992. “Primary Structure of the Aequorea Victoria Green-Fluorescent Protein.” *Gene* 111 (2): 229–33. doi:10.1016/0378-1119(92)90691-H.

Pries, F., J. Kingma, G. H. Krooshof, C. M. Jeronimus-Stratingh, A. P. Bruins, and D. B. Janssen. 1995. “Histidine 289 Is Essential for Hydrolysis of the Alkyl-Enzyme Intermediate of Haloalkane Dehalogenase.” *Journal of Biological Chemistry*. doi:10.1074/jbc.270.18.10405.

Prowazek, S Von. 1914. “Zur Kenntnis Der Giemsafarbung von Standpunkt Der Zytologie.” *Z Mikrosk* 31: 1–16.

Puchner, E. M., J. M. Walter, R. Kasper, B. Huang, and W. A. Lim. 2013. “Counting Molecules in Single Organelles with Superresolution Microscopy Allows Tracking of the Endosome Maturation Trajectory.” *Proceedings of the National Academy of Sciences* 110 (40): 16015–20. doi:10.1073/pnas.1309676110.

Rayleigh. 1879. “XXXI. *Investigations in Optics, with Special Reference to the Spectroscope*.” *Philosophical Magazine Series 5* 8 (49): 261–74. doi:10.1080/14786447908639684.

Ricci, Maria Aurelia, Carlo Manzo, Maria Filomena Garcia-Parajo, M. Lakadamyali, and Maria Pia Cosma. 2015. “Chromatin Fibers Are Formed by Heterogeneous Groups of Nucleosomes in Vivo.” *Cell* 160 (6): 1145–58. doi:10.1016/j.cell.2015.01.054.

Roosild, Tarmo P., Samantha Castronovo, and Senyon Choe. 2006. “Structure of Anti-FLAG M2 Fab Domain and Its Use in the Stabilization of Engineered Membrane Proteins.” *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* 62 (9): 835–39. doi:10.1107/S1744309106029125.

Rosenbloom, A. B., S.-H. Lee, M. To, A. Lee, J. Y. Shin, and C. Bustamante. 2014. “Optimized Two-Color Super Resolution Imaging of Drp1 during Mitochondrial Fission with a Slow-Switching Dronpa Variant.” *Proceedings of the National Academy of Sciences* 111 (36): 13093–98. doi:10.1073/pnas.1320044111.

Rothbauer, Ulrich, Kourosh Zolghadr, Sergei Tillib, Danny Nowak, Lothar Schermelleh,

Anja Gahl, Natalija Backmann, et al. 2006. “Targeting and Tracing Antigens in Live Cells with Fluorescent Nanobodies.” *Nature Methods* 3 (11): 887–89. doi:10.1038/nmeth953.

Rothemund, Paul W. K. 2006. “Folding DNA to Create Nanoscale Shapes and Patterns.” *Nature* 440 (7082): 297–302. doi:10.1038/nature04586.

Rust, Michael J, M. Bates, and Xiaowei Zhuang. 2006. “Sub-Diffraction-Limit Imaging by Stochastic Optical Reconstruction Microscopy (STORM).” *Nature Methods* 3 (10): 793–96. doi:10.1038/nmeth929.

Saerens, Dirk, Jörg Kinne, Eugène Bosmans, Ulrich Wernery, Serge Muyldermans, and Katja Conrath. 2004. “Single Domain Antibodies Derived from Dromedary Lymph Node and Peripheral Blood Lymphocytes Sensing Conformational Variants of Prostate-Specific Antigen.” *Journal of Biological Chemistry* 279 (50): 51965–72. doi:10.1074/jbc.M409292200.

Sage, Daniel, Hagai Kirshner, Thomas Pengo, Nico Stuurman, Junhong Min, Suliana Manley, and Michael Unser. 2015. “Quantitative Evaluation of Software Packages for Single-Molecule Localization Microscopy.” *Nature Methods* 12 (8): 717–24. doi:10.1038/nmeth.3442.

Sakin, Volkan, Janina Hanne, Jessica Dunder, Maria Anders-Össwein, Vibor Laketa, Ivana Nikić, Hans Georg Kräusslich, Edward A. Lemke, and Barbara Müller. 2017. “A Versatile Tool for Live-Cell Imaging and Super-Resolution Nanoscopy Studies of HIV-1 Env Distribution and Mobility.” *Cell Chemical Biology* 24 (5): 635–645.e5. doi:10.1016/j.chembiol.2017.04.007.

Sako, Yasushi, Shigeru Minoghchi, and Toshio Yanagida. 2000. “Single-Molecule Imaging of EGFR Signalling on the Surface of Living Cells.” *Nature Cell Biology* 2 (3): 168–72. doi:10.1038/35004044.

Salic, Adrian, and Timothy J Mitchison. 2008. “A Chemical Method for Fast and Sensitive Detection of DNA Synthesis in Vivo.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (7): 2415–20. doi:10.1073/pnas.0712168105.

Schermelleh, L., P. M. Carlton, S. Haase, L. Shao, L. Winoto, P. Kner, B. Burke, et al. 2008. “Subdiffraction Multicolor Imaging of the Nuclear Periphery with 3D

Structured Illumination Microscopy." *Science* 320 (5881): 1332–36. doi:10.1126/science.1156947.

Schnaible, V, and M Przybylski. 1999. "Identification of Fluorescein-5 '-Isothiocyanate-Modification Sites in Proteins by Electrospray-Ionization Mass Spectrometry." *Bioconjugate Chemistry* 10 (5): 861–66. doi:papers3://publication/uuid/E894E623-E09B-4225-A56E-FD367547FE07.

Schnitzbauer, Joerg, Maximilian T. Strauss, Thomas Schlichthaerle, Florian Schueder, and Ralf Jungmann. 2017. "Super-Resolution Microscopy with DNA-PAINT." *Nature Protocols* 12 (6). Nature Publishing Group: 1198–1228. doi:10.1038/nprot.2017.024.

Schoffelen, Sanne, Mark H L Lambermon, Mark B Van Eldijk, and Jan C M Van Hest. 2008. "Site-Specific Modification of Candida Antarctica Lipase B via Residue-Specific Incorporation of a Non-Canonical Amino Acid Site-Specific Modification of Candida Antarctica Lipase B via Residue-Specific Incorporation of a Non-Canonical Amino Acid." *Bioconjugate Chemistry*, 1–10. doi:10.1021/bc800019v.

Sengupta, Prabuddha, Tijana Jovanovic-Talisman, Dunja Skoko, Malte Renz, Sarah L Veatch, and J. Lippincott-Schwartz. 2011. "Probing Protein Heterogeneity in the Plasma Membrane Using PALM and Pair Correlation Analysis." *Nature Methods* 8 (11): 969–75. doi:10.1038/nmeth.1704.

Sengupta, Prabuddha, and J. Lippincott-Schwartz. 2012. "Quantitative Analysis of Photoactivated Localization Microscopy (PALM) Datasets Using Pair-Correlation Analysis." *BioEssays* 34 (5): 396–405. doi:10.1002/bies.201200022.

Shaner, Nathan C, Robert E Campbell, Paul A Steinbach, Ben N G Giepmans, Amy E Palmer, and Roger Y Tsien. 2004. "Improved Monomeric Red, Orange and Yellow Fluorescent Proteins Derived from Discosoma Sp. Red Fluorescent Protein." *Nature Biotechnology* 22 (12): 1567–72. doi:10.1038/nbt1037.

Sharonov, A., and R. M. Hochstrasser. 2006. "Wide-Field Subdiffraction Imaging by Accumulated Binding of Diffusing Probes." *Proceedings of the National Academy of Sciences* 103 (50): 18911–16. doi:10.1073/pnas.0609643104.

Shroff, Hari, Catherine G Galbraith, James a Galbraith, Helen White, Jennifer Gillette, Scott Olenych, Michael W Davidson, and E. Betzig. 2007. "Dual-Color

Superresolution Imaging of Genetically." *Proceedings of the National Academy of Sciences of the United States of America* 104 (51): 20308–13. doi:10.1073/pnas.0710517105.

Singh, Pranveer, Likhesh Sharma, S. Rajendra Kulothungan, Bharat V. Adkar, Ravindra Singh Prajapati, P. Shaik Syed Ali, Beena Krishnan, and Raghavan Varadarajan. 2013. "Effect of Signal Peptide on Stability and Folding of Escherichia Coli Thioredoxin." *PLoS ONE* 8 (5). doi:10.1371/journal.pone.0063442.

Sletta, H., A. Tøndervik, S. Hakvåg, T. E. Vee Aune, A. Nedal, R. Aune, G. Evensen, S. Valla, T. E. Ellingsen, and T. Brautaset. 2007. "The Presence of N-Terminal Secretion Signal Sequences Leads to Strong Stimulation of the Total Expression Levels of Three Tested Medically Important Proteins during High-Cell-Density Cultivations of Escherichia Coli." *Applied and Environmental Microbiology* 73 (3): 906–12. doi:10.1128/AEM.01804-06.

Small, Alex, and Shane Stahlheber. 2014. "Corrigendum: Fluorophore Localization Algorithms for Super-Resolution Microscopy." *Nature Methods* 11 (9): 971–971. doi:10.1038/nmeth0914-971a.

Smith, G. 1985. "Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface." *Science* 228 (4705): 1315–17. doi:10.1126/science.4001944.

So, Min kyung, Hequan Yao, and Jianghong Rao. 2008. "HaloTag Protein-Mediated Specific Labeling of Living Cells with Quantum Dots." *Biochemical and Biophysical Research Communications* 374 (3): 419–23. doi:10.1016/j.bbrc.2008.07.004.

Specht, ChristianG, Ignacio Izeddin, PamelaC Rodriguez, Mohamed ElBeheiry, Philippe Rostaing, Xavier Darzacq, Maxime Dahan, and Antoine Triller. 2013. "Quantitative Nanoscopy of Inhibitory Synapses: Counting Gephyrin Molecules and Receptor Binding Sites." *Neuron* 79 (2). Elsevier Inc.: 308–21. doi:10.1016/j.neuron.2013.05.013.

Stagge, Franziska, Gyuzel Y. Mitronova, Vladimir N. Belov, Christian A. Wurm, and Stefan Jakobs. 2013. "Snap-, CLIP- and Halo-Tag Labelling of Budding Yeast Cells." *PLoS ONE* 8 (10): 1–9. doi:10.1371/journal.pone.0078745.

Stasevich, Timothy J., Yoko Hayashi-Takanaka, Yuko Sato, Kazumitsu Maehara, Yasuyuki Ohkawa, Kumiko Sakata-Sogawa, Makio Tokunaga, et al. 2014. “Regulation of RNA Polymerase II Activation by Histone Acetylation in Single Living Cells.” *Nature* 516 (7530). Nature Publishing Group: 272–75. doi:10.1038/nature13714.

Steiner, Daniel, Patrik Forrer, Michael T Stumpp, and Andreas Plückthun. 2006. “Signal Sequences Directing Cotranslational Translocation Expand the Range of Proteins Amenable to Phage Display.” *Nature Biotechnology* 24 (7): 823–31. doi:10.1038/nbt1218.

Subach, Fedor V., George H Patterson, Suliana Manley, Jennifer M Gillette, J. Lippincott-Schwartz, and Vladislav V Verkhusha. 2009. “Erratum: Photoactivatable MCherry for High-Resolution Two-Color Fluorescence Microscopy.” *Nature Methods* 6 (4): 311–311. doi:10.1038/nmeth0409-311.

Subach, Fedor V., George H Patterson, Malte Renz, J. Lippincott-Schwartz, and Vladislav V Verkhusha. 2010. “For Two-Color Super-Resolution SptPALM of Live Cells.” *Cell*, no. 10: 12651–56. doi:10.1021/ja100906g.

Szymborska, Anna, A. de Marco, Nathalie Daigle, Volker C Cordes, John A G Briggs, and Jan Ellenberg. 2013. “Nuclear Pore Scaffold Structure Analyzed by Super-Resolution Microscopy and Particle Averaging.” *Science* 341 (6146): 655–58. doi:10.1126/science.1240672.

Takayama, Yuko, Hiroshi Sato, Shigeaki Saitoh, Yuki Ogiyama, Fumie Masuda, and Kohta Takahashi. 2008. “Biphasic Incorporation of Centromeric Histone CENP-A in Fission Yeast.” Edited by Kerry Bloom. *Molecular Biology of the Cell* 19 (2). The American Society for Cell Biology: 682–90. doi:10.1091/mbc.E07-05-0504.

Testa, Ilaria, Christian A. Wurm, Rebecca Medda, Ellen Rothermel, Claas Von Middendorf, Jonas Fölling, Stefan Jakobs, Andreas Schönle, S. W. Hell, and Christian Eggeling. 2010. “Multicolor Fluorescence Nanoscopy in Fixed and Living Cells by Exciting Conventional Fluorophores with a Single Wavelength.” *Biophysical Journal* 99 (8): 2686–94. doi:10.1016/j.bpj.2010.08.012.

Thorley, Jennifer A., Jeremy Pike, and Joshua Z. Rappoport. 2014. “Super-Resolution Microscopy: A Comparison of Commercially Available Options.” *Fluorescence*

Microscopy: Super-Resolution and Other Novel Techniques, 199–212. doi:10.1016/B978-0-12-409513-7.00014-2.

Turkowyd, Bartosz, David Virant, and Ulrike Endesfelder. 2016. “From Single Molecules to Life: Microscopy at the Nanoscale.” *Analytical and Bioanalytical Chemistry* 408 (25). Analytical and Bioanalytical Chemistry: 6885–6911. doi:10.1007/s00216-016-9781-8.

Uno, Shin-nosuke, Mako Kamiya, Toshitada Yoshihara, Ko Sugawara, Kohki Okabe, Mehmet C. Tarhan, Hiroyuki Fujita, et al. 2014. “A Spontaneously Blinking Fluorophore Based on Intramolecular Spirocyclization for Live-Cell Super-Resolution Imaging.” *Nature Chemistry* 6 (8). Nature Publishing Group: 681–89. doi:10.1038/nchem.2002.

Veatch, Sarah L., Benjamin B. Machta, Sarah A. Shelby, Ethan N. Chiang, David A. Holowka, and Barbara A. Baird. 2012. “Correlation Functions Quantify Super-Resolution Images and Estimate Apparent Clustering Due to over-Counting.” *PLoS ONE* 7 (2). doi:10.1371/journal.pone.0031457.

Vira, Shaleen, Elena Mekhedov, Glen Humphrey, and Paul S. Blank. 2010. “Fluorescent-Labeled Antibodies: Balancing Functionality and Degree of Labeling.” *Analytical Biochemistry* 402 (2). Elsevier Inc.: 146–50. doi:10.1016/j.ab.2010.03.036.

Virant, David, Bartosz Turkowyd, Alexander Balinovic, and Ulrike Endesfelder. 2017. “Combining Primed Photoconversion and UV-Photoactivation for Aberration-Free, Live-Cell Compliant Multi-Color Single-Molecule Localization Microscopy Imaging.” *International Journal of Molecular Sciences* 18 (7). doi:10.3390/ijms18071524.

Vreja, Ingrid C., Ivana Nikić, Fabian Göttfert, M. Bates, Katharina Kröhnert, Tiago F. Outeiro, S. W. Hell, Edward A. Lemke, and Silvio O. Rizzoli. 2015. “Super-Resolution Microscopy of Clickable Amino Acids Reveals the Effects of Fluorescent Protein Tagging on Protein Assemblies.” *ACS Nano* 9 (11): 11034–41. doi:10.1021/acsnano.5b04434.

Wang, Jiangyun, Jianming Xie, and Peter G. Schultz. 2006. “A Genetically Encoded Fluorescent Amino Acid.” *Journal of the American Chemical Society* 128 (27):

8738–39. doi:10.1021/ja062666k.

Wang, Lei, Jianming Xie, and Peter G. Schultz. 2006. “Expanding the Genetic Code.” *Annual Review of Biophysics and Biomolecular Structure* 35 (1): 225–49. doi:10.1146/annurev.biophys.35.101105.121507.

Williamson, David J, Dylan M Owen, Jérémie Rossy, Astrid Magenau, Matthias Wehrmann, J Justin Gooding, and Katharina Gaus. 2011. “Pre-Existing Clusters of the Adaptor Lat Do Not Participate in Early T Cell Signaling Events.” *Nature Immunology* 12 (7). Nature Publishing Group: 655–62. doi:10.1038/ni.2049.

Wilmes, Stephan, Markus Staufenbiel, Domenik Liße, Christian P. Richter, Oliver Beutel, K. B. Busch, Samuel T. Hess, and Jacob Piehler. 2012. “Triple-Color Super-Resolution Imaging of Live Cells: Resolving Submicroscopic Receptor Organization in the Plasma Membrane.” *Angewandte Chemie - International Edition* 51 (20): 4868–71. doi:10.1002/anie.201200853.

Wolter, S., Anna Löschberger, Thorge Holm, Sarah Aufmkolk, M.-C. Dabauvalle, Sebastian Van De Linde, Markus Sauer, and S. Van De Linde. 2012. “RapidSTORM : Accurate , Fast Open-Source Software for Localization Microscopy Orcae : Online Resource for Community Annotation of Eukaryotes.” *Nature Methods* 9 (11): 1040–41. doi:10.1038/nmeth.2224.

Wolter, S., M. Schüttpelz, M. Tscherepanow, S. Van De Linde, M. Heilemann, and M. Sauer. 2010. “Real-Time Computation of Subdiffraction-Resolution Fluorescence Images.” *Journal of Microscopy* 237 (1): 12–22. doi:10.1111/j.1365-2818.2009.03287.x.

Wombacher, Richard, Meike Heidbreder, Sebastian Van De Linde, Michael P. Sheetz, Mike Heilemann, Virginia W. Cornish, and Markus Sauer. 2010. “Live-Cell Super-Resolution Imaging with Trimethoprim Conjugates.” *Nature Methods* 7 (9). Nature Publishing Group: 717–19. doi:10.1038/nmeth.1489.

Xu, Ke, Hazen P Babcock, and Xiaowei Zhuang. 2012. “Dual-Objective STORM Reveals Three-Dimensional Filament Organization in the Actin Cytoskeleton.” *Nature Methods* 9 (2): 185–88. doi:10.1038/nmeth.1841.

Yao, Jie, and Robert Tjian. 2011. “Sub-Nuclear Compartmentalization of Core Promoter Factors and Target Genes.” *Cell Cycle* 10 (15): 2405–6.

doi:10.4161/cc.10.15.16199.

Zanacchi, Francesca Cella, Carlo Manzo, Angel S Alvarez, Nathan D Derr, Maria F Garcia-Parajo, and M. Lakadamyali. 2017. “A DNA Origami Platform for Quantifying Protein Copy Number in Super-Resolution.” *Nature Methods* 14 (8): 789–92. doi:10.1038/nmeth.4342.

Zessin, Patrick J M, Kieran Finan, and M. Heilemann. 2012. “Super-Resolution Fluorescence Imaging of Chromosomal DNA.” *Journal of Structural Biology* 177 (2). Elsevier Inc.: 344–48. doi:10.1016/j.jsb.2011.12.015.

Zhang, Yan, Min Kyung So, Andreas M. Loening, Hequan Yao, Sanjiv S. Gambhir, and Jianghong Rao. 2006. “HaloTag Protein-Mediated Site-Specific Conjugation of Bioluminescent Proteins to Quantum Dots.” *Angewandte Chemie - International Edition* 45 (30): 4936–40. doi:10.1002/anie.200601197.

Zhang, Zhengyang, Samuel J Kenny, Margaret Hauser, Wan Li, and Ke Xu. 2015. “Ultrahigh-Throughput Single-Molecule Spectroscopy and Spectrally Resolved Super-Resolution Microscopy.” *Nature Methods* 12 (10): 935–38. doi:10.1038/nmeth.3528.

Zhao, Z. W., R. Roy, J. C. M. Gebhardt, D. M. Suter, A. R. Chapman, and X. S. Xie. 2014. “Spatial Organization of RNA Polymerase II inside a Mammalian Cell Nucleus Revealed by Reflected Light-Sheet Superresolution Microscopy.” *Proceedings of the National Academy of Sciences* 111 (2): 681–86. doi:10.1073/pnas.1318496111.